

Struktur, Funktion und Allergenität
Bet v 1-homologer Proteine:
Das Allergen Gly m 4 und das Enzym Norcoclaurin-Synthase

Dissertation

zur Erlangung des Doktorgrades
der Fakultät für Biologie, Chemie und Geowissenschaften
der Universität Bayreuth

vorgelegt von

Dipl.-Biochem.

Hanna Berkner

aus Erlangen

Bayreuth, 2009

Die vorliegende Arbeit wurde von November 2005 bis Februar 2009 am Lehrstuhl für Struktur und Chemie der Biopolymere unter der Leitung von Prof. Dr. Paul Rösch angefertigt.

Vollständiger Abdruck der von der Fakultät für Biologie, Chemie und Geowissenschaften der Universität Bayreuth genehmigten Dissertation zur Erlangung des akademischen Grades eines Doktors der Naturwissenschaften (Dr. rer. nat.)

Promotionsgesuch eingereicht am: 05.02.2009

Tag des wissenschaftlichen Kolloquiums: 25.06.2009

Prüfungsausschuss:

Prof. Dr. Paul Rösch (Erster Gutachter)

Prof. Dr. Matthias Ullmann (Zweiter Gutachter)

Prof. Dr. Andreas Fery (Vorsitzender)

Prof. Dr. Carlo Unverzagt

Inhaltsverzeichnis

Inhaltsverzeichnis.....	I
Zusammenfassung.....	II
Summary	IV
1 Einleitung	1
1.1 Pharmakologisch wirksame Sekundärmetabolite aus Pflanzen	1
1.1.1 Kurze Geschichte der Pflanzenmedizin	1
1.1.2 Die Benzylisochinolinalkaloide	1
1.1.3 Der Biosyntheseweg der Benzylisochinolinalkaloide.....	2
1.1.4 Die Norcoclaurin-Synthase als Schlüsselenzym der BIA-Synthese	4
1.2 Pollen- und Nahrungsmittelallergene aus Pflanzen	7
1.2.1 Geschichte und Epidemiologie der Allergien	7
1.2.2 Die allergische Reaktion	8
1.2.3 Klassifizierung pflanzlicher Allergene.....	10
1.2.3.1 Klassifizierung aus taxonomischer Sicht	10
1.2.3.2 Klassifizierung in Proteinfamilien	11
1.2.4 Kreuzallergien und ihre molekulare Basis	13
1.2.5 Sojaallergie und –allergene	14
1.2.6 Die Familie der Bet v 1-homologen Allergene	15
1.2.6.1 Physiologische Funktion	16
1.2.6.2 Strukturelle Eigenschaften	17
1.2.6.3 Gly m 4 als Mitglied der Bet v 1-Familie	18
1.2.6.4 Norcoclaurin-Synthase als Mitglied der Bet v 1-Superfamilie	18
2 Ziele der Arbeit	20
3 Zusammenfassung und Diskussion der Ergebnisse	21
3.1 Gly m 4 - Ein klassisches Mitglied der Bet v 1-Familie.....	21
3.2 Norcoclaurin-Synthase – Ein neues Mitglied der Bet v 1-Superfamilie.....	22
3.2.1 Produktion des rekombinanten Proteins.....	22
3.2.2 Der Oligomerisierungszustand und die methodischen Konsequenzen	24
3.2.3 Das semi-experimentelle Homologiemodell.....	26
3.2.4 Vergleich des Homologiemodells mit der Kristallstruktur	27
3.2.5 Substratbindung unter Berücksichtigung der Komplexstruktur.....	28
3.2.6 Katalytischer Mechanismus	31
3.2.7 NCS-Aktivität und die physiologische Funktion der Bet v 1-Familie.....	33
4 Ausblick: <i>Epitope grafting</i>	35
5 Abkürzungsverzeichnis	37
6 Literaturverzeichnis.....	38
7 Publikationsliste	47
7.1 Einzelarbeit A.....	47
7.2 Einzelarbeit B.....	47
7.3 Einzelarbeit C.....	48
8 Einzelarbeiten.....	49
9 Danksagung.....	82
10 Erklärung.....	85

Zusammenfassung

In Nord- und Mitteleuropa wird ein beträchtlicher Teil der Pollen- und Nahrungsmittelallergien bei Erwachsenen durch allergene Proteine aus der Bet v 1-Familie ausgelöst, die nach dem Hauptbirkenpollenallergen benannt ist. Die Bet v 1-Familie wird in die Klasse 10 der *pathogenesis-related proteins* (PR10-Proteine) eingeordnet. Die PR10-Proteine spielen eine Rolle im „Immunsystem“ der Pflanze, wobei die genaue physiologische Funktion dieser Proteine bis heute nicht bestimmt werden konnte. Die strukturelle Untersuchung der Mitglieder der Bet v 1-Familie kann nicht nur neue Erkenntnisse bezüglich ihrer physiologischen Funktion erbringen, sondern stellt auch eine wichtige Voraussetzung für die Herstellung hypoallergener Proteinvarianten zur Verwendung in der Immuntherapie dar. Bisher konnten die Strukturen dreier verschiedener Bet v 1-Allergene gelöst werden, die alle die typische Bet v 1-Faltungstopologie aufweisen: ein siebensträngiges, antiparalleles β -Faltblatt und zwei kurze V-förmig angeordnete α -Helices, die zusammen mit einer langen C-terminalen α -Helix einen hydrophoben Hohlraum umschließen.

Die in dieser Arbeit mittels magnetischer Kernspinresonanz-(NMR-)Spektroskopie im Detail bestimmte Struktur des Bet v 1-Homologs Gly m 4 aus der als Nahrungsmittelzusatz verwendeten Sojabohne gehört ebenfalls diesem Faltungstyp an. In einigen strukturellen Einzelheiten aber zeigt Gly m 4 größere Ähnlichkeit zu drei PR10-Proteinen aus der gelben Lupine als zu Bet v 1. Die Ergänzung der gewonnenen strukturellen Informationen durch immunologische Daten ermöglichte die Lokalisierung von vier potentiellen kreuzreaktiven IgE-Epitopen auf der Oberfläche von Gly m 4, welche die molekulare Grundlage für Kreuzallergien auf Birkenpollen und Sojaprodukte darstellen.

Auch die enzymatisch aktiven (S)-Norcoclaurin-Synthasen (NCS) zeigen Sequenzähnlichkeiten mit Bet v 1, jedoch in wesentlich geringerem Maße als die „klassischen“ Bet v 1-Allergene. Im Rahmen dieser Arbeit konnte durch Kombination von Circular dichroismus- und NMR-Spektroskopie mit der Methode der Homologiemodellierung ein semi-experimentelles Strukturmodell der NCS aus der Alkaloid-produzierenden Pflanze *Thalictrum flavum* konstruiert werden. Anhand dieses Modells konnte gezeigt werden, dass auch die NCS die typische Bet v 1-Faltungstopologie aufweist, wobei jedoch die C-terminale α -Helix in drei Abschnitte unterteilt ist. Erst kürzlich wurde das Modell weitgehend durch die Kristallstruktur des Proteins bestätigt. Die NCS ist somit ein echtes Mitglied der Bet v 1-Superfamilie, in der sie im Hinblick auf ihre bekannte physiologische Funktion eine Ausnahme darstellt. Sie

katalysiert einen wichtigen Schritt in der Synthese pharmakologisch aktiver Sekundärmetabolite, die Kondensation von 4-Hydroxyphenylacetaldehyd und Dopamin zu (S)-Norcocclaurin. Anhand von NMR-Titrationsexperimenten mit Substrat bzw. Substrat-analogen und Untersuchung des Oligomerisierungszustands in An- und Abwesenheit der Substrate konnten erste Einblicke in das aktive Zentrum des Enzyms und dessen Reaktionsmechanismus gewonnen werden. Aktivitätstests mit Bet v 1 zeigten, dass das Hauptbirkenpollenallergen hingegen nicht in der Lage ist, die NCS-Reaktion zu katalysieren. Bezüglich des Nutzens der NCS für die Erforschung kreuzreaktiver Epitope stellt das Enzym aufgrund seiner im Zuge dieser Arbeit ermittelten strukturellen und immunologischen Eigenschaften einen interessanten Kandidaten für das sogenannte *Epitope grafting* dar. Ziel dieser Methode ist es, eine bestimmte Bindungsstelle eines Proteins auf ein anderes Protein zu übertragen. In diesem Fall soll mit dem Ziel der detaillierten Charakterisierung der Antikörper-Bindungsstelle ein potentiell IgE-Epitop von Bet v 1 auf die strukturell sehr ähnliche, aber kaum kreuzreaktive NCS übertragen werden.

Summary

In Northern and Central Europe, a considerable percentage of pollen and food allergies in adults is caused by allergenic proteins of the Bet v 1-family. This protein family whose name is derived from the major birch pollen allergen belongs to class 10 of the pathogenesis-related proteins (PR10 proteins). These proteins play a role in the ‘immune system’ of the plant, but their precise physiological function has not been determined yet. Structural studies of members of the Bet v 1-family will not only provide new information regarding their physiological function, but are also an important prerequisite for the production of hypo-allergenic protein variants for immunotherapy. So far, the structures of three different Bet v 1-allergens have been solved. All of them show the typical Bet v 1-fold: a seven-stranded, antiparallel β -sheet and two short V-shaped α -helices which in combination with a long C-terminal α -helix form a hydrophobic cavity in the protein.

During this work, the three-dimensional structure of the Bet v 1-homologue Gly m 4 from soybean was determined in detail by means of nuclear magnetic resonance (NMR) spectroscopy. This allergen is frequently found in different kinds of processed foods as soy is used as a food additive. The protein shows the same overall fold as Bet v 1, but certain structural details are more similar to three PR10 proteins from yellow lupine. The correlation of the structural information with immunological data paved the way for the localization of four putative cross-reactive IgE epitopes on the allergen’s surface. These epitopes represent the molecular basis for cross-reactivities on birch pollen and soy products.

The enzymatically active (S)-norcoclaurine synthases (NCS) show sequence similarities to Bet v 1 as well, however in a considerably lower extent compared to the ‘classical’ Bet v 1-allergens. In this work, we combined circular dichroism and NMR spectroscopy with homology modeling to build a semi-experimental model of NCS from the alkaloid-producing plant *Thalictrum flavum*. This model shows that NCS adopts the typical Bet v 1-fold, although the C-terminal α -helix is divided into three parts. Very recently, the crystal structure of the protein has been solved, confirming the model in major parts. NCS is thus a true member of the Bet v 1-superfamily. With respect to its well-characterized physiological function, the enzyme represents an exception among the Bet v 1-homologous PR10 proteins. It catalyzes an important step during the synthesis of pharmacologically active secondary metabolites, namely the condensation of 4-hydroxyphenylacetaldehyde and dopamine to (S)-norcoclaurine. By means of NMR titration experiments with the substrates or analogues, respectively, and determination of the oligomerization state in absence or presence of the

substrates, first insights into the active site of the enzyme and its catalytic mechanism could be provided. Enzyme activity assays with Bet v 1 showed that the major birch pollen allergen is not able to catalyze the NCS reaction. The structural and immunological properties of NCS determined during this work make the enzyme an interesting candidate for epitope grafting. The general purpose of this method is to transfer binding sites between different proteins. Here we tried to transfer a putative IgE epitope of Bet v 1 to the structurally similar but nearly non-cross-reactive NCS aiming to characterize the respective epitope in detail.

1 Einleitung

1.1 Pharmakologisch wirksame Sekundärmetabolite aus Pflanzen

1.1.1 Kurze Geschichte der Pflanzenmedizin

Die Nutzung von Pflanzen zu medizinischen Zwecken hat in der Geschichte der Menschheit eine sehr lange Tradition. Reste verschiedener Heilpflanzen, die bei Ausgrabungen entdeckt wurden, zeugen davon, dass bereits in der Steinzeit von ihnen Gebrauch gemacht wurde (Chevallier, 2000). Um 1000 v. Chr. entstand in Indien die Ayurveda-Medizin, die zum Beispiel die Wurzeln der Alkaloid-produzierenden Pflanze *Rauwolfia serpentina* einsetzte (Buchanan *et al.*, 2000). Die Nutzung des Schlafmohns *Papaver somniferum* (Abb. 1) zu medizinischen Zwecken kann bis ca. 1400 v. Chr. zurückverfolgt werden. Während in der griechisch-römischen Kultur das aus der Fruchtkapsel der Pflanze gewonnene Opium vermischt mit Wein und weiteren Zutaten unter dem Namen Theriak als Gegengift bei diversen Vergiftungen eingesetzt wurde, wussten die Benediktinermönche um 800 bereits Extrakte der Pflanze als Schmerz- und Betäubungsmittel zu verwenden (Buchanan *et al.*, 2000; Grabley & Thiericke, 2000). Die Entdeckung der chemischen Substanzen, die für die medizinische Wirkung der Pflanzen verantwortlich sind, nahm zu Beginn des 19. Jahrhunderts ihren Anfang. Wiederum war es der Schlafmohn, aus dem der deutsche Pharmakologe Friedrich Sertürner 1806 das analgetisch wirksame Morphin isolierte (Buchanan *et al.*, 2000). Damit war der Grundstein für die Erforschung und kommerzielle Nutzung pharmakologisch wirksamer Sekundärmetabolite aus Pflanzen gelegt.



Abb. 1: *Papaver somniferum*
(www.imagines-plantarum.de)

1.1.2 Die Benzylisochinolinalkaloide

Das auch heute noch häufig als Schmerzmittel eingesetzte Morphin gehört zur ca. 2500 Substanzen umfassenden Gruppe der Benzylisochinolinalkaloide (BIA), die wiederum eine Untergruppe der ca. 12000 pflanzlichen Alkaloide darstellt (Wink, 1999a; Liscombe & Facchini, 2008). Alkaloide sind in der Pflanzenwelt weit verbreitet. 40 % aller Pflanzenfamilien umfassen mindestens eine Alkaloid-produzierende Art (Harborne *et al.*, 1999). Die

Verbreitung der BIA im Pflanzenreich hingegen beschränkt sich mit wenigen Ausnahmen auf die zur Klasse der zweikeimblättrigen Bedecktsamer gehörigen Unterklassen *Magnoliidae* und *Ranunculidae* (Wink, 1999a). In der Pflanze erfüllen diese Sekundärmetabolite aufgrund ihrer häufig toxischen oder antimikrobiellen Wirkung die Aufgabe des Schutzes vor Fressfeinden oder mikrobiellem Befall.

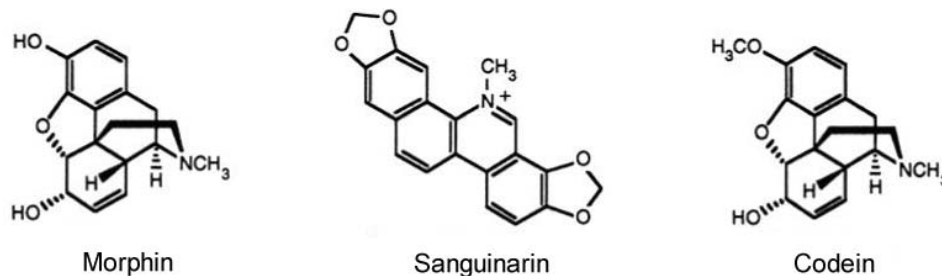


Abb. 2: Strukturen verschiedener pharmakologisch wirksamer Benzylisochinolinalkaloide

Vom Menschen werden aufgrund ihrer vielfältigen pharmakologischen Aktivitäten neben Morphin noch weitere Verbindungen dieser Substanzklasse in der modernen Medizin eingesetzt (Abb. 2). So ist das strukturell eng mit Morphin verwandte Codein häufig als Antitussivum in Hustensäften enthalten. Das weniger bekannte Sanguinarin wird aufgrund seiner antibakteriellen, entzündungshemmenden Wirkung häufig in der Mundhygiene eingesetzt (Harborne *et al.*, 1999).

1.1.3 Der Biosyntheseweg der Benzylisochinolinalkaloide

Die Biosynthesewege der verschiedenen pflanzlichen BIA sind relativ gut erforscht. Der erste Alkaloidsyntheseweg, in dem man alle beteiligten Enzyme lückenlos identifizieren und isolieren konnte, war der des antimikrobiell wirksamen Berberins (Buchanan *et al.*, 2000). Die ersten Schritte der Synthese sind für alle BIA gleich. Zunächst wird die Aminosäure Tyrosin zu Dopamin bzw. 4-Hydroxyphenylacetaldehyd (4-HPAA) umgesetzt. Im nächsten Schritt entsteht durch die stereospezifische Kondensation dieser beiden Substanzen (S)-Norcoclaurin, das erste Intermediat mit Benzylisochinolin-Grundgerüst. (S)-Norcoclaurin wird anschließend in mehreren Schritten zu (S)-Reticulin umgesetzt (Abb. 3). Während der Syntheseweg der Bisbenzylisochinolin-Alkaloide bereits während der Umwandlung von (S)-Norcoclaurin zu (S)-Reticulin abzweigt, stellt (S)-Reticulin einen wichtigen Verzweigungspunkt für die Synthese weiterer nach ihrer chemischen Struktur klassifizierter BIA dar. So trennen sich hier die Synthesewege der Morphin und Codein umfassenden

Morphinan-Alkaloide, der Benzophenanthridin-Alkaloide, zu denen Sanguinarin zählt, und der Protoberberin-Alkaloide (Ziegler & Facchini, 2008; Abb. 3).

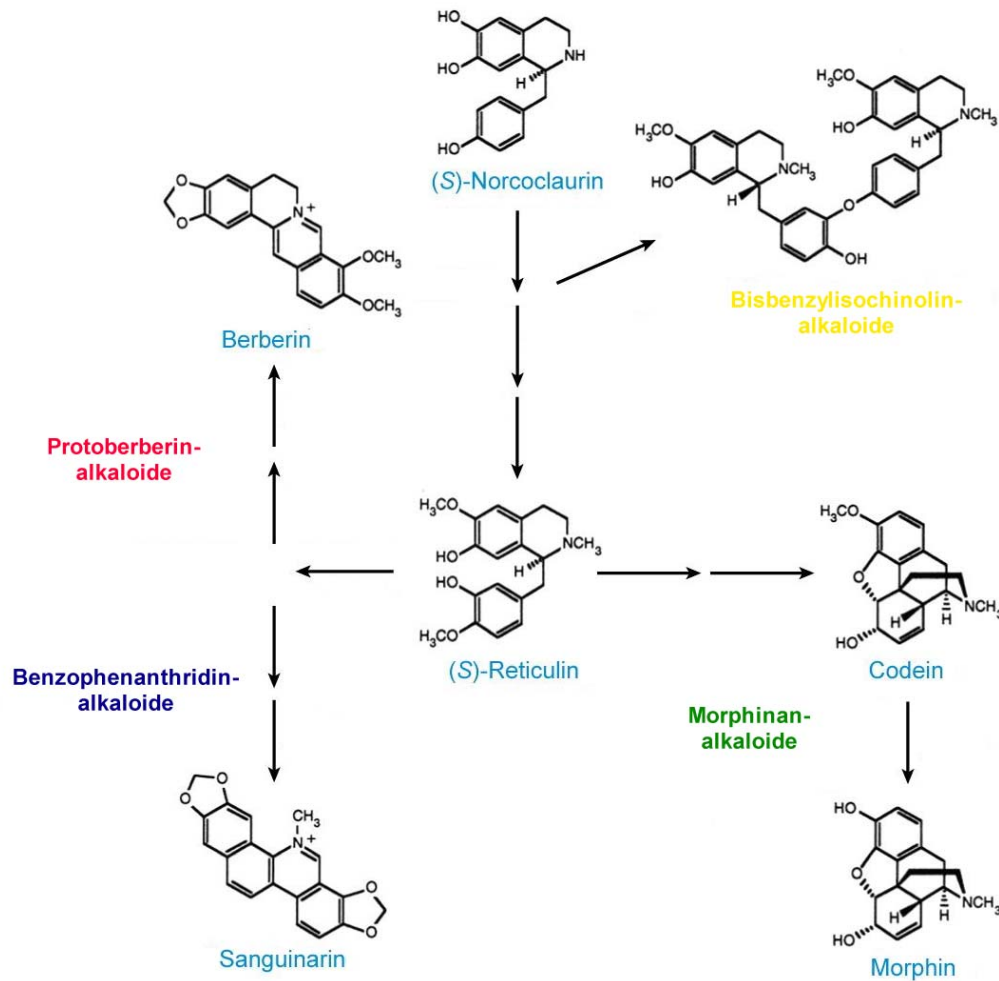


Abb. 3: Biosynthese der Benzylisochinolinalkaloide

Schematische Darstellung des Synthesewegs der Bisbenzylisochinolinalkaloide, der Protoberberin-Alkaloide, der Benzophenanthridin-Alkaloide und der Morphinan-Alkaloide ausgehend von ihrem gemeinsamen Vorläufer (S)-Norcoclaurin (nach Ziegler & Facchini, 2008).

Da viele medizinisch nutzbare BIA wegen ihrer strukturellen Komplexität gar nicht oder nur unter hohem Aufwand chemisch synthetisierbar sind, stellt die Isolierung aus der Pflanze oder aus Pflanzenzellkulturen häufig den einzigen Weg zur Gewinnung der entsprechenden BIA dar. Detaillierte Kenntnisse über die biosynthetischen Enzyme und deren Regulation sind daher unabdingbare Voraussetzung für die Optimierung der Ausbeute der gewünschten pharmakologisch aktiven Substanz (Wink, 1999b).

1.1.4 Die Norcoclaurin-Synthase als Schlüsselenzym der BIA-Synthese

Die (S)-Norcoclaurin-Synthase (NCS) kann insofern als Schlüsselenzym der BIA-Synthese betrachtet werden, als dieses Enzym die Reaktion katalysiert, die zum ersten Zwischenprodukt mit Benzyloisochinolin-Grundgerüst führt. Heute weiß man, dass es sich bei diesem gemeinsamen Vorläufer aller BIA um (S)-Norcoclaurin handelt. Zu Beginn des 20. Jahrhunderts wurde jedoch zunächst aufgrund theoretischer Überlegungen das tetrahydroxylierte (S)-Norlaudanosolin als Reaktionsprodukt der Kondensation von Dopamin und 3,4-Dihydroxyphenylacetaldehyd (3,4-DHPAA) als gemeinsamer Vorläufer der BIA postuliert (Winterstein & Trier, 1910; Robinson, 1917). Bis in die achtziger Jahre hinein wurde diskutiert, inzwischen auch basierend auf experimentellen Erkenntnissen (Battersby & Binks, 1960; Wilson & Coscia, 1975), ob neben Dopamin tatsächlich 3,4-DHPAA oder aber 3,4-Dihydroxyphenylpyruvat umgesetzt wird, was zur Entstehung von Norlaudanosolin-1-carboxylsäure führen würde (Abb. 4A). Diese wiederum könnte durch Decarboxylierung zu (S)-Norlaudanosolin umgesetzt werden.

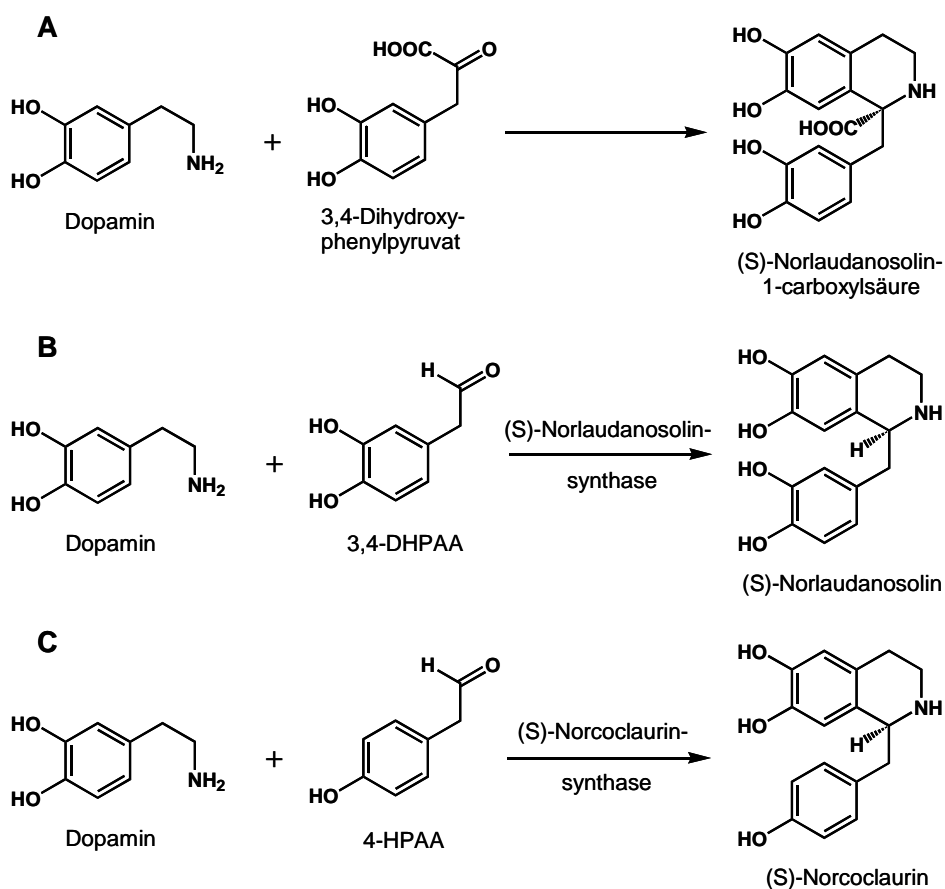


Abb. 4: Die Bildung des Vorläufers aller pflanzlichen BIA

Historische Vorschläge von A) Battersby & Binks (1960) bzw. B) Rueffer *et al.* (1981) und C) die tatsächlich von der NCS katalysierte Reaktion zur Bildung des gemeinsamen BIA-Vorläufers (S)-Norcoclaurin (Stadler & Zenk, 1990)

Im Jahre 1981 wurde jedoch mittels Radioisotopenmarkierung der potentiellen Substrate bestätigt, dass mehrere BIA-produzierende Pflanzen ein Enzym besitzen, welches die Kondensation von 3,4-DHPAA und Dopamin zu (S)-Norlaudanosolin katalysiert. 3,4-Dihydroxyphenylpyruvat hingegen wurde nicht als Substrat akzeptiert. Das Enzym wurde daraufhin (S)-Norlaudanosolin-Synthase genannt (Rueffer *et al.*, 1981; Abb. 4B). Weitere Experimente, bei denen unter anderem die Metabolisierung von radioisotopenmarkiertem (S)-Norcoclaurin in BIA-produzierenden Pflanzen verfolgt wurde, zeigten jedoch, dass anstelle des tetrahydroxylierten (S)-Norlaudanosolin nur das trihydroxylierte (S)-Norcoclaurin als Intermediat der BIA-Biosynthese in der Pflanze vorkommt (Stadler *et al.*, 1987; Loeffler *et al.*, 1987; Stadler *et al.*, 1989). Somit wurde das Enzym, welches demzufolge in der Pflanze die Kondensation von Dopamin und 4-HPAA zu (S)-Norcoclaurin katalysiert, in NCS umbenannt (Stadler & Zenk, 1990; Abb. 4C).

Die Isolierung der NCS aus verschiedenen Pflanzenarten wie dem Schlafmohn *Papaver somniferum* oder der weniger bekannten gelben Wiesenraute *Thalictrum flavum* war die erste Voraussetzung für die Untersuchung ihrer enzymatischen Eigenschaften (Samanani & Facchini, 2001; Samanani & Facchini, 2002). Einen weiteren wichtigen Schritt auf dem Weg zur umfassenden Charakterisierung des Enzyms stellte die Identifizierung des entsprechenden Gens und dessen Klonierung dar, die wiederum die heterologe Produktion des rekombinanten Proteins in bakteriellen Expressionssystemen ermöglichte (Samanani *et al.*, 2004; Liscombe *et al.*, 2005; Minami *et al.*, 2007). Die aus *Thalictrum flavum* isolierte NCS (*Tf*-NCS) war 2004 das erste Enzym dieser Art, welches als rekombinantes Protein hergestellt werden konnte (Samanani *et al.*, 2004). Sowohl die aus verschiedenen Pflanzen isolierten als auch die rekombinanten Proteine wurden bereits auf ihre enzymatischen Eigenschaften wie Substratkinetiken, pH- und Temperaturoptimum untersucht (Samanani & Facchini, 2001; Samanani & Facchini, 2002; Samanani *et al.*, 2004; Liscombe *et al.*, 2005; Minami *et al.*, 2007). Auch zum Mechanismus der katalysierten Reaktion wurden wiederum an der *Tf*-NCS Untersuchungen durchgeführt. Die NCS katalysiert eine asymmetrische Pictet-Spengler-Reaktion. Bei diesem Reaktionstyp handelt es sich um eine elektrophile Zyklisierung zwischen einem Iminiumion und einem aromatischen Ring. Im Falle der NCS-Reaktion wird durch Reaktion der 4-HPAA-Carbonylgruppe mit der Aminogruppe des Dopamins das Iminiumion gebildet, welches anschließend am C5 des von Dopamin stammenden Benzolrings angreift. Anhand von Experimenten mit verschiedenen Substratanaloga wurde erst kürzlich der in Abb. 5 dargestellte Reaktionsmechanismus vorgeschlagen (Luk *et al.*, 2007).

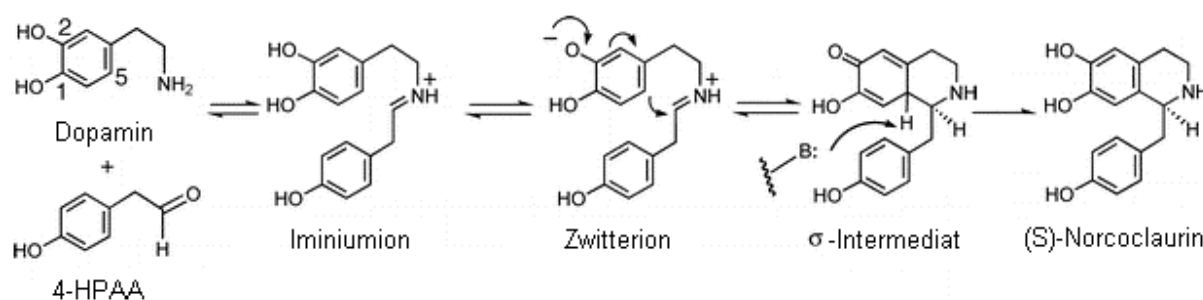


Abb. 5: **Vorgeschlagener Mechanismus der NCS-Reaktion** (nach Luk *et al.*, 2007)

Während die NCS in Bezug auf ihre enzymatischen Eigenschaften und den Mechanismus der katalysierten Reaktion schon detailliert untersucht wurde, lagen zu Beginn dieser Arbeit noch keinerlei strukturelle Informationen über das Enzym vor, die jedoch zur Charakterisierung des aktiven Zentrums unabdingbar sind. Eine genaue Kenntnis der an Substratbindung und Katalyse beteiligten Aminosäuren ist etwa. in Hinblick auf die Optimierung der enzymatischen Herstellung pharmakologischer Wirkstoffe von großem Interesse.

1.2 Pollen- und Nahrungsmittelallergene aus Pflanzen

1.2.1 Geschichte und Epidemiologie der Allergien

Pflanzen haben nicht nur als Lieferant pharmakologisch aktiver Sekundärmetabolite eine medizinische Bedeutung für den Menschen. Sie sind auch Quelle von Proteinen, die beim Menschen eine allergische Reaktion hervorrufen können. Allergien werden oft als ein Phänomen der jüngeren Vergangenheit betrachtet. Es existieren jedoch einige geschichtliche Dokumente, die das Auftreten allergischer Reaktionen bereits vor unserer Zeitrechnung belegen. So enthält ein auf ca. 1600 v. Chr. datiertes Papyrus eine detaillierte Beschreibung der Symptome des allergischen Asthmas. Ein prominenter Fall einer Nahrungsmittelallergie mit schweren Folgen, allerdings nicht medizinischer sondern politischer Art, ist aus der Regierungszeit Richards des Dritten überliefert. Der König schrieb die Symptome einer durch Erdbeeren ausgelösten Allergie einem vermeintlichen Giftanschlag zu und ließ den als Täter beschuldigten unbequemen Lord William Hastings hinrichten. Der tatsächlichen Ursache allergischer Reaktionen kam man jedoch erst im 19. Jahrhundert auf die Spur, als der englische Arzt Charles Blackley im Selbstversuch nachweisen konnte, dass Pollenkörner die Symptome des Heuschnupfen hervorrufen können (Schumacher, 1995). Der Begriff „Allergie“ wurde erst zu Beginn des 20. Jahrhunderts vom damals berühmten österreichischen Kinderarzt Clemens von Pirquet im weitesten Sinne als „veränderte Reaktionsfähigkeit“ des Organismus auf eine Substanz, gegen die vorher eine Sensibilisierung stattgefunden hat, definiert (von Pirquet, 1906; Schumacher, 1995). Seitdem hat sich das Fachgebiet der Allergologie zu einem wichtigen und anerkannten Forschungszweig entwickelt.

Die Beobachtung, dass die Prävalenz allergischer Erkrankungen vor allem in der industrialisierten Welt in den letzten Jahren teils rapide angestiegen ist (Kay, 2001), kann aber nicht allein dem gestiegenen „Allergie-Bewusstsein“ zugeschrieben werden. Mehrere Faktoren werden für diesen Anstieg verantwortlich gemacht. So besagt zum Beispiel die „Hygiene-Hypothese“, dass Infektionen im frühen Kindesalter durch Stimulation bestimmter Zellen des Immunsystems dem späteren Auftreten von Allergien entgegenwirken können (Wills-Karp *et al.*, 2001; Garn & Renz, 2007). Dies würde erklären, dass in Entwicklungsländern mit schlechteren hygienischen Verhältnissen eine geringere Prävalenz allergischer Erkrankungen zu beobachten ist (Romagnani, 2004). Auch die genetische Prädisposition spielt für das Allergierisiko eine Rolle. Menschen, die eine solche Prädisposition aufweisen, bezeichnet man als atopisch. Durch Zwillings- und Familienstudien konnten bereits einige Gene identifiziert werden, die möglicherweise zur Atopie beitragen

können, da sie für Proteine kodieren, die an der allergischen Reaktion beteiligt sind. Dazu gehören das auf dem Chromosom 11 lokalisierte Gen für die β -Untereinheit des Immunoglobulin E-Rezeptors und die auf dem Chromosom 5 befindlichen Gene für eine Reihe von Cytokinen (Janeway *et al.*, 2002). Als weitere Faktoren, die sich auf die Prävalenz von Allergien auswirken, werden die Allergenexposition an sich und die Schadstoffbelastung diskutiert (Heppt & Bachert, 1998). So konnte z.B. nachgewiesen werden, dass Dieselruß eine Verstärkung der Produktion von proinflammatorischen Molekülen hervorrufen kann, die in der Signalübertragung während der allergischen Reaktion eine wichtige Rolle spielen (Ring *et al.*, 2001; Takano *et al.*, 2007).

1.2.2 Die allergische Reaktion

Die immunologischen Hypersensibilitätsreaktionen werden nach Coombs und Gell je nach Mechanismus in vier verschiedene Typen unterteilt (Gell & Coombs, 1963). Neben der Sofortreaktion vom Typ I, die durch Immunoglobuline der Klasse E (IgE) vermittelt wird, wird außerdem die zytotoxische IgG-vermittelte Reaktion des Typ II, die Immunkomplexreaktion des Typ III und die verzögerte zellvermittelte Typ IV-Reaktion unterschieden. Reaktionen des Typ II können z.B. durch Medikamentenallergien gegen Penicillin ausgelöst werden. Zu den Typ III-Reaktionen gehört die sogenannte Serumkrankheit, die z.B. nach Injektion von zu therapeutischen Zwecken hergestelltem Pferdeserum auftreten kann. Ein Beispiel für eine Typ IV-Reaktion ist die allergische Kontaktdermatitis, die unter anderem durch Kontakt mit nickelhaltigen Materialien ausgelöst werden kann (Schumacher, 1995; Garner, 2004). Da die klassischen allergischen Symptome wie allergische Rhinitis, allergisches Asthma oder systemische Anaphylaxie, die durch Pollen- oder Pollen-assoziierte Nahrungsmittelallergene hervorgerufen werden, dem Reaktionstyp I angehören, wird hier nur dieser Mechanismus ausführlicher erläutert (Abb. 6). Der erste Schritt der Typ I-Reaktion ist die Sensibilisierung, die zur Produktion von allergenspezifischem IgE führt. Dabei wird das Antigen bzw. das Allergen zunächst von dendritischen Zellen aufgenommen und in Peptide gespalten. Diese Peptide werden dann gebunden an den Haupthistokompatibilitäts-Komplex des Typ II (MHC II) auf der Zelloberfläche präsentiert. Der MHC II-Peptid-Komplex wird vom T-Zell-Rezeptor naiver CD4-T-Zellen erkannt, was über eine Signaltransduktionskaskade zur Produktion bestimmter Cytokine führt (Janeway *et al.*, 2002). Die Cytokine wiederum regulieren die Differenzierung der T-Helferzellen (Th-Zellen) zu Th1- oder Th2-Zellen. So fördert z.B. Interleukin 4 (IL-4)

die Bildung von Th2-Zellen, während IL-12 die Bildung von Th1-Zellen stimuliert (Romagnani, 2006). Die Differenzierung ist jedoch nicht nur von den ausgeschütteten Cytokinen abhängig, sondern auch von den spezifischen Eigenschaften des Antigens, von seiner Dosis und von seinem Präsentationsweg (Janeway *et al.*, 2002). Handelt es sich bei dem Antigen um ein Allergen, so wird bei atopischen Individuen die Differenzierung zu Th2-Zellen begünstigt. Zwei verschiedene von diesen Th2-Zellen ausgehende Signale bewirken dann in B-Zellen den Isotypwechsel der dort produzierten Immunoglobuline von M nach E. Zum einen ist dies die Ausschüttung der Interleukine IL-4 oder IL-13, zum anderen die CD40-vermittelte Wechselwirkung der entsprechenden T- und B-Zellen (Miescher & Vogel, 2002; Davies & O'Hehir, 2008). Die von den B-Zellen sezernierten allergenspezifischen IgE-Moleküle binden anschließend an den hochaffinen FcεRI-Rezeptor, der sich auf der Oberfläche von Mastzellen befindet (Turner & Kinet, 1999).

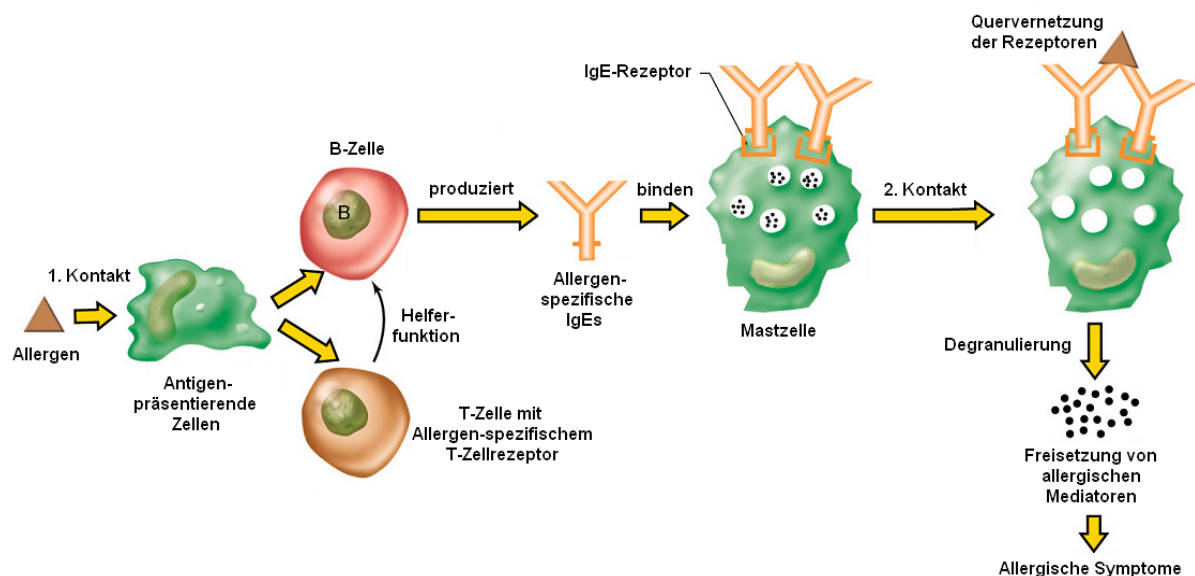


Abb. 6: Die allergische Reaktion

Schematische Darstellung der Sensibilisierung und allergischen Reaktion des Typ I (nach Madigan *et al.*, 2003, S.943)

Beim zweiten Kontakt mit dem Allergen führt die Erkennung der Epitope durch IgE-Moleküle zu einer Quervernetzung der Rezeptoren (Gould & Sutton, 2008). Dies ist das Signal zur Degranulierung der Mastzellen und somit zur Ausschüttung einer Reihe von proinflammatorischen Mediatoren wie z.B. Serotonin und Histamin. Gleichzeitig wird die Synthese von Eicosanoidhormonen wie Prostaglandinen oder Leukotrienen gefördert. Diese Substanzen verursachen dann die verschiedenen Symptome der allergischen Entzündungsreaktion, die von verstärkter Schleimsekretion und Juckreiz in den Augen bis hin zum lebensbedrohlichen anaphylaktischen Schock reichen (Venarske & deShazo, 2003; Gould *et al.*, 2003).

1.2.3 Klassifizierung pflanzlicher Allergene

Eine zentrale Frage der Allergenforschung lautet: Was macht ein Protein zum Allergen? Diese Frage kann bis heute nicht zufriedenstellend beantwortet werden. Durch die Identifikation, Klonierung und biochemische, immunologische und strukturelle Untersuchung einer Vielzahl von Allergenen werden jedoch immer mehr Informationen zur Beantwortung dieser Frage gewonnen. Eine Klassifizierung der Allergene nach bestimmten Eigenschaften verschafft eine Übersicht über Gemeinsamkeiten und Unterschiede dieser Moleküle. So können sie beispielsweise nach ihrer Quelle grob in Proteine tierischen und pflanzlichen Ursprungs unterteilt werden. Allergene aus Pflanzen spielen für den Menschen zum einen als in Pollenkörnern durch die Luft verbreitete inhalative Allergene eine Rolle (Mohapatra *et al.*, 2008; Esch, 2008; Swoboda *et al.*, 2008). Zum anderen treten sie als Nahrungsmittelallergene in Obst, Gemüse und in weiteren aus Pflanzen hergestellten Produkten auf (Burks & Ballmer-Weber, 2006). Dabei wird zwischen Nahrungsmittelallergenen unterschieden, die selbst zur Sensibilisierung fähig sind, und solchen, bei denen die Sensibilisierung durch ein strukturell verwandtes inhalatives Allergen erfolgt. Letztere bilden die Gruppe der Pollen-assoziierten Nahrungsmittelallergene (Breiteneder & Mills, 2005).

1.2.3.1 Klassifizierung aus taxonomischer Sicht

Eine Möglichkeit zur Klassifizierung von Pflanzenallergenen ist die Gruppierung nach ihrer botanischen Herkunft, die sich auch in der 1986 von der World Health Organization eingeführten allgemeingültigen Allergennomenklatur widerspiegelt (Marsh *et al.*, 1986). Danach werden Allergene mit den ersten drei Buchstaben der Gattung, gefolgt vom ersten Buchstaben der Art der Allergenquelle und einer chronologisch nach dem Zeitpunkt der Isolierung gerichteten laufenden Nummer bezeichnet. Wichtige Baumpollenallergene gehören zum Beispiel hauptsächlich drei verschiedenen Ordnungen an, und zwar den *Fagales*, den *Scrophulariales* und den *Coniferales*. Diese Klassifizierung bringt gleichzeitig eine geographische Einordnung mit sich. Die *Fagales* mit einheimischen Bäumen wie Birke, Erle, Hasel, Hainbuche, Eiche und Kastanie sind neben Nord- und Mitteleuropa auch in Nordwest-Afrika, Ostasien und auf dem amerikanischen Kontinent zu finden. Im Mittelmeerraum spielt der Olivenbaum als Mitglied der *Scrophulariales* eine wichtige Rolle als Pollenallergenquelle, wobei er auch in Südafrika, Australien und Amerika heimisch ist. Ein weiteres Mitglied dieser Ordnung, die Esche, ist in Mitteleuropa und den USA für Pollenallergien verantwortlich. Die in Hinsicht auf enthaltene Allergene bedeutendsten Bäume der Ordnung *Coniferales* sind

Zypresse, Zeder und Wacholder, die ebenfalls im Mittelmeerraum, aber auch in Australien, Neuseeland, Amerika, China und Indien wachsen (Mothes *et al.*, 2004). Eine weitere Pflanzenfamilie mit großer allergologischer Relevanz ist die der Gräser, der *Poaceae*. Zu ihr zählen ungefähr ein Viertel aller Pflanzenspezies, die Pollenallergene enthalten (Radauer & Breiteneder, 2006). Auch in Pollen der Familie *Asteraceae* sind Allergene enthalten. Die Gattung *Ambrosia* ist ein Beispiel für die ungewollte Einschleppung dieser ursprünglich in Nordamerika heimischen Pflanze in Europa, die eine stark verstärkte Prävalenz von Allergien gegen *Ambrosia*-Allergene nach sich zog (Oberhuber *et al.*, 2008).

1.2.3.2 Klassifizierung in Proteinfamilien

Die Allergenforschung hat in den letzten zwei Jahrzehnten große Fortschritte auf dem Gebiet der molekularen Charakterisierung von Allergenen erzielt. Viele immunologisch relevante Allergene konnten als rekombinantes Protein gewonnen werden und standen somit für biochemische und strukturelle Untersuchungen zur Verfügung. Die Sequenzen und teils auch weitergehende biochemische Informationen sind größtenteils in allgemein zugänglichen Datenbanken wie die der Organisation Allergome (<http://www.allergome.org>) oder die des europäischen Netzwerkes Protall (<http://www.ifr.ac.uk/protall>) hinterlegt (Schein *et al.*, 2007). Strukturelle Informationen zu bekannten Allergenen können über die Structural Database of Allergenic Proteins (SDAP, <http://fermi.utmb.edu/SDAP>) erhalten werden, die über 40 PDB-Strukturen allergener Proteine bereithält (Schein *et al.*, 2007; Chapman *et al.*, 2007). All diese Informationen ermöglichen eine Einordnung der Allergene in Proteinfamilien. Mittlerweile gibt es auch eine Datenbank, über die man direkt die Einordnung des Allergens in die entsprechende Proteinfamilie erhält (Allfam, <http://www.meduniwien.ac.at/allergens/allfam>; Radauer *et al.*, 2008a). Während Pflanzenallergene taxonomisch gesehen über eine Vielzahl verschiedener Spezies verteilt sind, können sie auf molekularer Ebene einer sehr beschränkten Anzahl an Proteinfamilien zugeordnet werden. So stammen alle bekannten Pollenallergene aus nur 29 von 7868 beschriebenen Proteinfamilien. Die pflanzlichen Nahrungsmittelallergene stellen sogar nur Mitglieder von 27 verschiedenen Proteinfamilien, wovon 10 Familien auch Pollenallergene enthalten (Radauer & Breiteneder, 2006). Ein Großteil der Pollenallergene zählt zur Familie der Expansine, der Profiline oder der Calcium-bindenden Proteine, während über 60 % aller pflanzlichen Nahrungsmittelallergene zu den Superfamilien der Prolamine oder der Cupine, zu den Profilinen, oder zu den Verwandten des Hauptbirkenpollenallergens Bet v 1 zählen (Radauer & Breiteneder, 2007).

Abb. 7 zeigt die dreidimensionalen Strukturen einiger allergener Repräsentanten dieser Proteinfamilien.

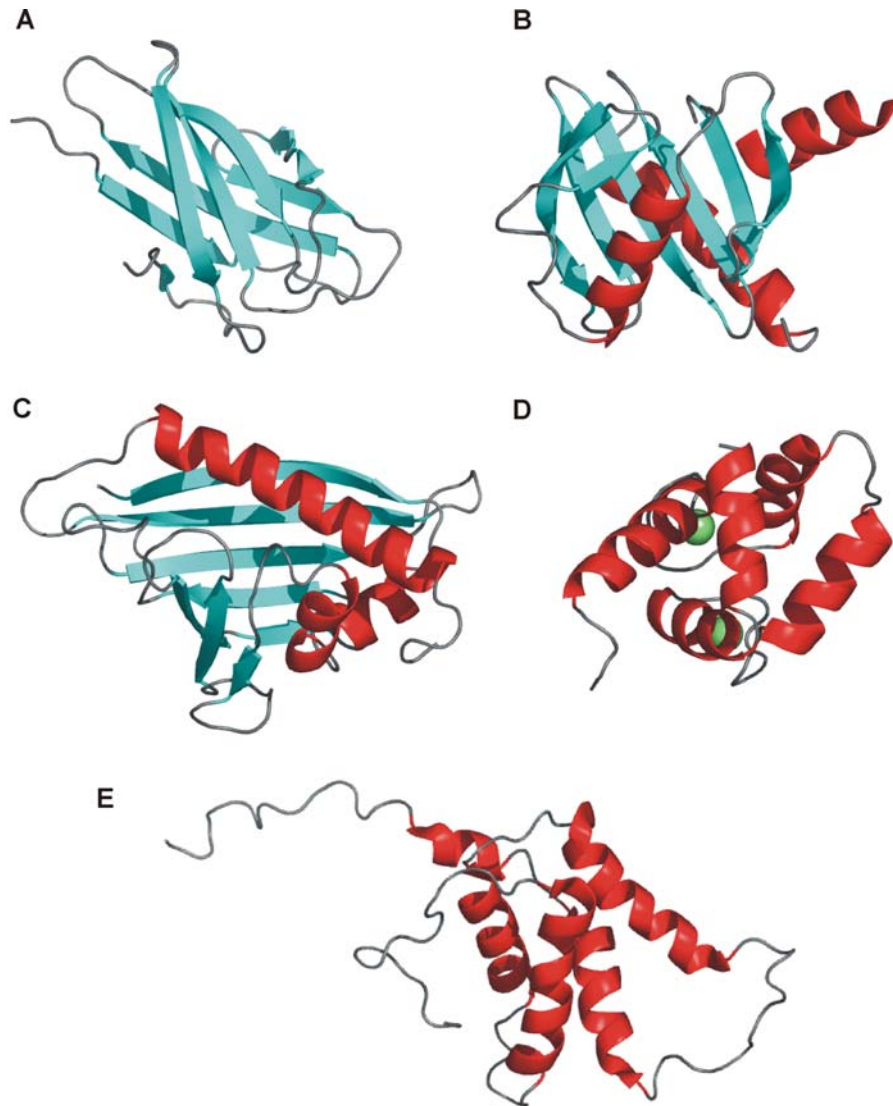


Abb. 7: Allergene Proteine aus verschiedenen Proteinfamilien

A) Das Expansin Phl p 3 (PDB-Code 2JNZ); B) das Profilin Bet v 2 (PDB-Code 1CQA); C) Pru av 1 aus der Bet v 1-Familie (PDB-Code 1E09); D) das Calcium-bindende Bet v 4 (PDB-Code 1H4B); das Prolamin Ara h 6 (PDB-Code 1W2Q).

Bei der Betrachtung der Strukturen fällt ins Auge, dass die Proteine unterschiedlichen Faltungstypen angehören. So bildet das Expansin und Graspollenallergen Phl p 3 eine reine β -Faltblattstruktur aus (Schweimer *et al.*, 2008), wohingegen das Calcium-bindende EF-Hand-Protein Bet v 4 aus Birkenpollen und das zur Familie der Prolamine gehörige 2S-Albumin Ara h 6 aus der Erdnuss zu den rein α -helikalen Proteinen gehören (Neudecker *et al.*, 2004; Lehmann *et al.*, 2006). Das ebenfalls in Birkenpollen enthaltene Profilin Bet v 2 und das Bet v 1-verwandte Nahrungsmittelallergen Pru av 1 zählen zu den gemischten α/β -Proteinen (Fedorov *et al.*, 1997; Neudecker *et al.*, 2001). Die Allergenität eines Proteins kann somit also nicht aufgrund eines bestimmten Faltungstyps vorausgesagt

werden. Dennoch spricht die geringe Zahl an Proteinfamilien mit allergenen Mitgliedern dafür, dass bestimmte molekulare Eigenschaften mitverantwortlich für die Allergenität der Proteine sind. Die Profilin beispielsweise erfüllen mit ihrer eher geringen Masse von etwa 14 kDa und ihrer guten Wasserlöslichkeit zwei wichtige Voraussetzungen für ein Pollenallergen, da sie somit in der Lunge leicht aus den Pollen herausgelöst werden und mit dem Immunsystem in Kontakt treten können. Nahrungsmittelallergene wiederum müssen zur Sensibilisierung den Bedingungen des Gastrointestinaltraktes standhalten. Diese Eigenschaft trifft auf Proteine der Prolamin- und der Cupin-Superfamilie mit ihrer Resistenz gegenüber Proteasen und ihrer Säure- und Thermostabilität zu. Proteine der Bet v 1-Superfamilie hingegen sind eher Proteolyse-anfällig und thermolabil und gehören deshalb zur Gruppe der Pollen-assoziierten Nahrungsmittelallergene (Mills & Shewry, 2004).

1.2.4 Kreuzallergien und ihre molekulare Basis

Während die dreidimensionale Struktur keinen Rückschluss auf die Allergenität eines Proteins erlaubt, spielt sie für das Phänomen der Kreuzallergien eine wichtige Rolle. Kreuzallergien basieren auf der Fähigkeit von allergen-spezifischen IgEs zur Bindung von Allergenen aus anderen Quellen, die zum sensibilisierenden Allergen homolog sind. Das sensibilisierende Allergen und das kreuzreaktive Allergen weisen also auf ihrer Oberfläche ähnliche Bindungsstellen bzw. IgE-Epitope auf. Da es sich bei IgE-Epitopen hauptsächlich um konformationelle Epitope handelt (Ferreira *et al.*, 1998), ist für eine solche Kreuzreaktivität sowohl ein gewisses Maß an zumindest lokaler Sequenzidentität als auch eine ähnliche dreidimensionale Struktur der betreffenden Allergene erforderlich.

Kreuzallergien können sowohl zwischen nah verwandten Spezies, die ausschließlich inhalative Allergene enthalten, wie den zur Ordnung *Fagales* zählenden Baumarten, als auch zwischen phylogenetisch weit entfernten Spezies wie Birke und Kiwi auftreten (Mothes *et al.*, 2004; Gall *et al.*, 1994). Letzteres Phänomen ist ein Beispiel für eine pollen-assoziierte Nahrungsmittelallergie.

Auf molekularer Ebene können neben Allergenen der Bet v 1-Superfamilie auch Profilin wie Bet v 2 (Abb. 7B) für Kreuzallergien verantwortlich sein. Diese Aktin-bindenden Proteine kommen nahezu in allen Eukaryonten vor. Während ca. 90 % aller Birkenpollenallergiker eine Sensibilisierung gegen Bet v 1 aufweisen, ist in ca. 20 % aller Pollenallergiker ein Profilin für die Sensibilisierung verantwortlich (Vieths *et al.*, 2002). Profilin mit Sequenzidentitäten zu Bet v 2 zwischen 71 und 83 % konnten in verschiedenen Obstsorten,

wie Kirsche, Birne, Apfel, Banane, Litschi, Ananas, aber auch in Gemüse, Gewürzen und Nüssen, wie Sellerie, Soja, Tomaten, Paprika, Erdnuss und Haselnuss, identifiziert werden (Vieths *et al.*, 2002). Weitere Allergene, die zumindest durch *in vitro*-Inhibitionstests als kreuzreaktiv eingestuft werden konnten, sind Homologe des Birkenpollenallergens Bet v 6, welches Phenylcumaran-Benzylether-Reduktase-Aktivität aufweist (Karamloo *et al.*, 2001), sowie Bet v 4-verwandte Polcalcine aus Kräutern und Gräsern (Hayek *et al.*, 1998; Verdino *et al.*, 2008). Auch zu den Prolaminen gehörige unspezifische Lipidtransferproteine wie Pru p 3 aus Pfirsich und Zea m 14 aus Mais zeigen Kreuzreaktivität (Pastorello *et al.*, 2000; Breiteneder & Mills, 2006). Kreuzreaktive IgEs spezifisch für bestimmte Glykosilierungen, sogenannte kreuzreaktive Kohlenhydrat-Determinanten, konnten zwar *in vitro* in einer erheblichen Anzahl von allergischen Patienten detektiert werden, haben aber eine sehr geringe klinische Relevanz (Altmann, 2007). Glykosylierte Allergene sind z.B. das Vicilin Ara h 1 aus der Erdnuss und Lyc e 2, eine β -Fructofuranosidase aus der Tomate (van Ree *et al.*, 2000; Westphal *et al.*, 2003).

1.2.5 Sojaallergie und –allergene

Sojabohnen sind ein traditionell in Asien und den USA weit verbreitetes Nahrungsmittel. In den letzten Jahren hat auch in Europa der Konsum von Produkten, die Soja enthalten, rapide zugenommen. Seit den 80er Jahren sind sowohl bei Kindern mit atopischer Dermatitis als auch bei Erwachsenen Fälle von Allergien gegen Sojaproteine dokumentiert. Da Sojabohnen nicht nur unverarbeitet als Gemüse sondern wegen ihrem hohen Proteingehalt häufig als Zusatz von prozessierten Nahrungsmitteln wie Fleisch- oder Backwaren verwendet werden, können Sojaproteine in diesen Nahrungsmitteln als versteckte Allergene auftreten. Trotz dieser weitreichenden Problematik gibt es bisher nur wenig Informationen über Sojaallergien (Ballmer-Weber & Vieths, 2008).

Bisher konnten mindestens 16 IgE-reaktive Proteine aus Sojabohnen identifiziert werden, wobei nur sechs davon bisher von der International Union of Immunological Societies offiziell als Allergene anerkannt wurden (<http://www.allergen.org>): die Hüllproteine Gly m 1 und Gly m 2, die bei der Verarbeitung von Soja respiratorische Allergien auslösen können (Gonzalez *et al.*, 1992), die Nahrungsmittelallergene Gly m 3 und Gly m 4, ein Profilin und ein Bet v 1-Homolog (Rihs *et al.*, 1999; Kleine-Tebbe *et al.*, 2002), sowie die zur Cupin-Superfamilie gehörigen Proteine Gly m 5 und Gly m 6 (Pedersen *et al.*, 2008). Studien mit europäischen Sojabohnenallergikern deuten darauf hin, dass es sich bei einem Großteil der

Sojaallergien in der erwachsenen Bevölkerung um Birkenpollen-assoziierte Nahrungsmittelallergien handelt, die hauptsächlich durch Kreuzreaktionen von Bet v 1-spezifischen IgEs mit dem homologen Sojaallergen Gly m 4 ausgelöst werden (Mittag *et al.*, 2004; Ballmer-Weber *et al.*, 2007). Diese Studien zeigten auch, dass bei Birkenpollen-assoziierten Nahrungsmittelallergien gegen Soja neben den typischen eher moderaten Symptomen wie dem durch Juckreiz im Mundraum charakterisierten oralen Allergie-Syndrom (OAS) auch schwerwiegende systemische Reaktionen wie ein anaphylaktischer Schock auftreten können.

1.2.6 Die Familie der Bet v 1-homologen Allergene

Die Familie der Bet v 1-homologen Allergene umfasst verschiedene Pollen- und Nahrungsmittelallergene aus Gemüse, Obst und Nüssen, die Sequenzhomologie zum Hauptallergen der Weißbirke (*Betula verrucosa*) Bet v 1 aufweisen (Tab. 1).

Tab. 1: Bisher identifizierte Allergene der Bet v 1-Familie (nach www.meduniwien.ac.at/allergens/allfam)

Name	Quelle	Expositionsroute	Aminosäuren ¹	Sequenzidentität mit Bet v 1A (%)	Uniprot-Eintrag
Aln g 1	Erle (<i>Alnus glutinosa</i>)	Inhalation	159	81	P38948
Api g 1	Sellerie (<i>Apium graveolens</i>)	Nahrung	153	41	P49372
Ara h 8	Erdnuß (<i>Arachis hypogaea</i>)	Nahrung	156	46	Q6VT83
Aspa o 17kD	Spargel (<i>Asparagus officinalis</i>)	Nahrung	157	54	Q05736
Bet p 1	Asiatische Weißbirke (<i>Betula platyphylla</i>)	Inhalation	159	96	Q9AYS2
Bet v 1	Weißbirke (<i>Betula verrucosa</i>)	Inhalation	159	100	P15494
Cap ch 17kD	Pepperoni (<i>Capsicum chinense</i>)	Nahrung	158	35	Q5DUH6
Car b 1	Hainbuche (<i>Carpinus betulus</i>)	Inhalation	159	72	P38949
Cas s 1	Kastanie (<i>Castanea sativa</i>)	Inhalation	159	67	Q93YH9
Cor a 1	Hasel (<i>Corylus avellana</i>)	Inhalation/ Nahrung	160	67	Q9SWR4
Cor he 1	Asiatische Hasel (<i>Corylus heterophylla</i>)	Inhalation	160	66	A8W7B6
Dau c 1	Karotte (<i>Daucus carota</i>)	Nahrung	153	37	O04298
Fag s 1	Buche (<i>Fagus sylvatica</i>)	Inhalation	159	68	Q9ZRU8
Fra a 1	Erdbeere (<i>Fragaria ananassa</i>)	Nahrung	159	53	Q3T923
Gly m 4	Sojabohne (<i>Glycine max</i>)	Nahrung	157	47	P26987
Mal d 1	Apfel (<i>Malus domestica</i>)	Nahrung	158	55	P43211
Pet c 1	Petersilie (<i>Petroselinum crispum</i>)	Nahrung	154	40	P19418
Pru ar 1	Aprikose (<i>Prunus armeniaca</i>)	Nahrung	159	59	O50001
Pru av 1	Kirsche (<i>Prunus avium</i>)	Nahrung	159	59	O24248
Pru p 1	Pfirsich (<i>Prunus persica</i>)	Nahrung	159	59	Q216V8
Pyr c 1	Birne (<i>Pyrus communis</i>)	Nahrung	158	57	O65200
Que a 1	Weiß-Eiche (<i>Quercus alba</i>)	Inhalation	n.d. ²	n.d. ²	P85126
Rub i 1	Himbeere (<i>Rubus idaeus</i>)	Nahrung	n.d. ²	n.d. ²	Q0Z8U9
Tar o 18kD	Löwenzahn (<i>Taraxacum officinale</i>)	n.d.	156	36	O49065
Vig r 1	Mungobohne (<i>Vigna radiata</i>)	Nahrung	154	44	Q2VU97

¹Anzahl der Aminosäuren unter der Annahme, dass bei allen Bet v 1-Homologen das N-terminale Methionin abgespalten wird

²Anzahl der Aminosäuren und Sequenzidentität konnte nicht bestimmt werden, da nur Fragmente der Proteinsequenzen hinterlegt sind

Das Hauptbirkenpollenallergen wurde 1983 erstmals gereinigt und immunologisch untersucht und daraufhin Bet v 1 genannt (Ipsen & Lowenstein, 1983). Sechs Jahre später konnte das für Bet v 1 kodierende Gen identifiziert und als eines der ersten Gene, die ein Allergen kodieren, kloniert werden (Breiteneder *et al.*, 1989). Die Bet v 1-Familie wird aufgrund von Sequenzhomologien zur Klasse 10 der *pathogenesis-related proteins* (PR-Proteine) gezählt. Die im Pflanzenreich weit verbreiteten PR-Proteine umfassen mittlerweile insgesamt 17 Klassen, von denen mindestens 9 Klassen Allergene enthalten (Hoffmann-Sommergruber, 2000; van Loon *et al.*, 2006). PR-Proteine können ebenso wie bestimmte bioaktive Sekundärmetabolite als Teil des „Immunsystems“ der Pflanze betrachtet werden. Meist wird ihre Expression durch mikrobielle Infektion, Verletzung der Pflanzensubstanz oder andere umweltbedingte Stressstimuli induziert, wobei auch entwicklungs- oder gewebespezifisch konstitutiv exprimierte PR-Proteine existieren (Ebner *et al.*, 2001). So konnte für das Hauptbirkenpollenallergen gezeigt werden, dass die Expression verschiedener Isoformen des Proteins durch Bakterien- und Pilzbefall induziert werden kann, während Bet v 1 selber konstitutiv exprimiert wird (Swoboda *et al.*, 1994; Hoffmann-Sommergruber *et al.*, 1997).

1.2.6.1 Physiologische Funktion

Die weite Verbreitung der PR10-Proteine im Pflanzenreich lässt auf eine wichtige physiologische Funktion dieser Proteine schließen. Deshalb finden sich zahlreiche Untersuchungen zur Funktionalität Bet v 1-homologer Proteine. So konnte für PR10-Proteine aus Ginseng Ribonuclease-Aktivität nachgewiesen werden, die *in-vitro*-Studien auch im Fall von Bet v 1 bestätigten, jedoch in sehr schwacher Ausprägung (Moiseyev *et al.*, 1997; Bufe *et al.*, 1996). Weiterhin legten strukturelle Ähnlichkeiten mit der START-Domäne des humanen Cholesterin-Transportproteins MLN64 eine Funktion als Steroidtransporter nahe (Tsujiyama & Hurley, 2000). Die daraufhin durchgeführten Bindungsstudien ergaben, dass sowohl Isoformen von Bet v 1 als auch das Hauptallergen der Kirsche, Pru av 1, zur Interaktion mit Phytohormonen aus der Klasse der Brassinosteroide fähig sind (Neudecker *et al.*, 2001; Markovic-Housley *et al.*, 2003). Zudem konnte gezeigt werden, dass Bet v 1 auch eine Reihe weiterer physiologisch relevanter Liganden wie Fettsäuren, Flavonoide, Cytokine oder Phospholipide binden kann (Mogensen *et al.*, 2002; Mogensen *et al.*, 2007). Ein als P-Schleife bekanntes innerhalb der Bet v 1-Familie konserviertes Sequenzmotiv wiederum deutet auf eine Funktion als Kinase oder Nukleotid-bindendes Protein hin (Saraste *et al.*,

1990). Trotz all dieser Ergebnisse konnte den PR10-Proteinen bis jetzt keine bestimmte physiologische Funktion zugeordnet werden.

1.2.6.2 Strukturelle Eigenschaften

Die Mitglieder der Bet v 1-Allergenfamilie sind kleine, globuläre Proteine mit Molekülmassen um 17 kDa, bestehen aus 153-160 Aminosäuren und weisen meist einen im sauren Bereich liegenden isoelektrischen Punkt auf (Vieths *et al.*, 2002). Außer der Abspaltung des NH₂-terminalen Methionins sind für diese Proteinfamilie keine weiteren post-translationalen Modifikationen bekannt (Breiteneder *et al.*, 1989; Swoboda *et al.*, 1995; Ipsen & Hansen, 1991; Schöning *et al.*, 1995; Luttkopf *et al.*, 2002). Alle Proteine der Bet v 1-Familie weisen an zu Aminosäuren 46-52 in Bet v 1 äquivalenter Sequenzposition das glycinreiche P-Schleifen-Motiv mit der Konsensussequenz G-X-G-G-X-G-T auf (Saraste *et al.*, 1990). Bisher konnten mit magnetischer Kernspinresonanz-(NMR-)Spektroskopie und Röntgenkristallographie die dreidimensionalen Strukturen von drei verschiedenen Vertretern der Bet v 1-Familie bestimmt werden. Die erste bekannte Struktur war die von Bet v 1 selbst (Gajhede *et al.*, 1996; Abb. 8).

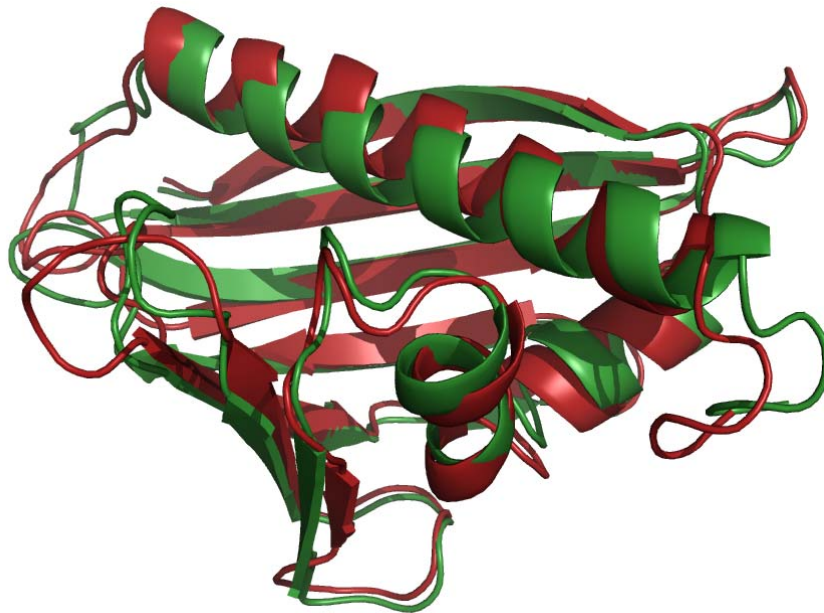


Abb. 8: Die kreuzreaktiven Allergene Bet v 1 und Pru av 1

Überlagerung der Kristallstruktur des Hauptbirkenpollenallergens Bet v 1 (grün, PDB-Code 1BV1) und des homologen Allergens Pru av 1 aus der Kirsche (ziegelrot, PDB-Code 1E09)

Sie besteht aus einem siebensträngigen antiparallelen β -Faltblatt und zwei kurzen in V-Form angeordneten α -Helices, die gemeinsam gegen eine lange C-terminale α -Helix packen. Diese Elemente umschließen einen hydrophoben Hohlraum, dem eine Rolle bei der Bindung

physiologischer Liganden zugeschrieben wird. Fünf Jahre später folgte die dreidimensionale Struktur von Pru av 1, das bei einer Sequenzidentität von 59 % eine exakt identische Faltung aufweist und damit auf eindrucksvolle Weise die molekulare Basis von Kreuzallergien veranschaulicht (Neudecker *et al.*, 2001; Abb. 8). Auch die 2005 gelöste Struktur des Sellerieallergens Api g 1 weist bei einer Sequenzidentität von 41 % die gleiche charakteristische Faltung auf wie Bet v 1 (Schirmer *et al.*, 2005).

1.2.6.3 Gly m 4 als Mitglied der Bet v 1-Familie

Das Sojabohnenallergen Gly m 4 ist mit 157 Aminosäuren, einem isoelektrischen Punkt von 4,4 (Vieths *et al.*, 2002) und einer Sequenzidentität zu Bet v 1 von 47 % ein typisches Bet v 1-Allergen. Das entsprechende Gen wurde 1992 identifiziert. Wegen der für PR-Proteine typischen Stressinduzierbarkeit und der entwicklungsspezifischen Regulation der Expression erhielt das entsprechende Protein zunächst die Bezeichnung SAM22 für *starvation-associated message* 22 (Crowell *et al.*, 1992). Zehn Jahre später konnte in einer Studie mit Birkenpollenallergikern gezeigt werden, dass SAM22 Birkenpollen-assoziierte Sojaallergien hervorrufen kann (Kleine-Tebbe *et al.*, 2002). Daraufhin wurde das Protein nach der gültigen Nomenklatur für Allergene in Gly m 4 umbenannt. Die Verbreitung von sojahaltigen Nahrungsmitteln in Europa und die teils schwerwiegenden Symptome der Allergie unterstreichen die Notwendigkeit der genauen Charakterisierung des auslösenden Moleküls.

1.2.6.4 Norcoclaurin-Synthase als Mitglied der Bet v 1-Superfamilie

In den letzten Jahren wurde die Klassifikation der Bet v 1-Familie durch die Einführung der Bet v 1-Superfamilie erweitert. Diese umfasst neben den PR10-Proteinen auch sogenannte *major latex-* und *ripening related proteins* (MLP und RRP) sowie Cytokinin-bindende Proteine (CSBP) und die NCS-Familie (Radauer & Breiteneder, 2007). Die Mitglieder dieser Proteinfamilien weisen bezogen auf Bet v 1 alle nur geringe Sequenzidentitäten von bis zu 25 % auf. Kürzlich zeigte jedoch die Bestimmung der Strukturen eines CSBP aus der Mungobohne *Vigna radiata* und eines MLP aus *Arabidopsis thaliana*, dass Vertreter dieser beiden Proteinfamilien ebenfalls entsprechend dem Bet v 1-Typ gefaltet sind (Pasternak *et al.*, 2006; PDB-Code 2I9Y). Die NCS und das PR10-Protein Hyp-1 aus Johanniskraut stellen eine Besonderheit innerhalb der Bet v 1-Superfamilie dar, da ihre enzymatische Funktion bekannt ist (Samanani & Facchini, 2001; Bais *et al.*, 2003; s. Abschnitt 1.1.4). Die Tf-NCS liegt mit 25 % Sequenzidentität mit Bet v 1 in der sogenannten „Twilight-Zone“ (Rost, 1999), d.h. eine

Vorhersage der Struktur allein aufgrund der Sequenzhomologie ist mit Skepsis zu betrachten. Desweiteren beinhaltet die Sequenz der *Tf*-NCS mit insgesamt 210 Aminosäuren im Vergleich zu den typischen Bet v 1-Allergenen sowohl einen längeren N-terminalen als auch einen kürzeren C-terminalen Überhang. Die ersten 19 Aminosäuren wurden als putatives Signalpeptid für den sekretorischen Transportweg identifiziert (Samanani *et al.*, 2004). Dies lässt vermuten, dass die *Tf*-NCS im Gegensatz zu den klassischen Bet v 1-Homologen kein cytosolisches Protein ist. Gerade diese Unterschiede machen die Beantwortung der Frage, ob auch die NCS die typische Bet v 1-Faltungstopologie aufweist, umso interessanter.

2 Ziele der Arbeit

Das Hauptziel dieser Arbeit war die detaillierte strukturelle Untersuchung des Bet v 1-homologen Sojabohnenallergens Gly m 4 und die strukturelle, funktionelle und immunologische Untersuchung des ebenfalls Bet v 1-homologen Enzyms NCS.

Die Bestimmung der Struktur des Sojabohnenallergens Gly m 4 mittels NMR-Spektroskopie sollte in Kombination mit immunologischen Daten Aufschluss über die Lokalisierung kreuzreaktiver IgE-Epitope auf der Proteinoberfläche geben, welche die molekulare Ursache von Birkenpollen-assoziierten Sojaallergien darstellen.

Das Enzym NCS aus der gelben Wiesenraute wurde bisher nur aufgrund seiner im Vergleich zu Gly m 4 eher geringen Sequenzhomologie zum Hauptbirkenpollenallergen mit der Bet v 1-Familie in Verbindung gebracht. Um das Enzym strukturellen Untersuchungen zugänglich zu machen war es nötig, die Produktion des rekombinanten Proteins zu optimieren. Darauf aufbauend sollte mit Hilfe von Circular dichroismus-(CD-) und NMR-Spektroskopie geklärt werden, ob die NCS die für die Bet v 1-Familie typische Faltungstopologie aufweist und somit ein echtes Mitglied der Bet v 1-Superfamilie darstellt. Immunologische Untersuchungen auf Kreuzreaktivität mit Bet v 1 und anderen homologen Allergenen sollten in Kombination mit den strukturellen Informationen zur Beantwortung der Frage beitragen, welche strukturellen Merkmale für die Allergenität einiger Mitglieder der Bet v 1-Superfamilie verantwortlich sind. Neue Hinweise auf die physiologische Funktion der Bet v 1-homologen Allergene sollte die Untersuchung von Bet v 1 auf die Fähigkeit zur Katalyse der NCS-Reaktion liefern.

Ein weiteres Ziel bestand in der Charakterisierung der enzymatischen Funktion der NCS, die eine Schlüsselrolle bei der Synthese pharmakologisch aktiver Sekundärmetabolite spielt. Dazu sollten mit Hilfe von NMR-Titrationsexperimenten mit den Substraten und Substratanaloga die Substratbindungsstellen ermittelt und das aktive Zentrum des Enzyms lokalisiert werden. In diesem Zusammenhang sollte auch der Oligomerisierungszustand des Proteins ermittelt werden.

3 Zusammenfassung und Diskussion der Ergebnisse

3.1 Gly m 4 - Ein klassisches Mitglied der Bet v 1-Familie

Das Allergen Gly m 4 aus der als Nahrungsmittel weltweit steigende Popularität genießenden Sojabohne gehört mit einer relativ hohen Sequenzidentität von 47 % bezogen auf Bet v 1 zu den „klassischen“ Bet v 1-Allergenen (Einzelarbeit A, Abb. 1). Anhand eines auf Bet v 1 und Pru av 1 basierenden Homologiemodells, das durch erste experimentelle Daten gestützt wurde, konnte bereits gezeigt werden, dass Gly m 4 die typische Bet v 1-Faltungstopologie aufweist (Neudecker *et al.*, 2003). Diese besteht aus einem siebensträngigen antiparallelen β -Faltblatt, welches in Kombination mit zwei kurzen, V-förmig angeordneten α -Helices gegen eine lange C-terminale α -Helix packt. Die genannten Elemente umschließen einen hydrophoben Hohlraum im Inneren des Proteins.

In dieser Arbeit konnte durch Einbezug von aus NOESY-Experimenten gewonnenen Abstandsbeschränkungen und von dynamischen Parametern eine rein experimentelle Struktur bestimmt werden, die sich im Detail vor allem in drei Punkten von den bereits bekannten Strukturen der Bet v 1-Allergene unterscheidet (Einzelarbeit A, Abb. 3). Erstens werden die Faltblattstränge $\beta 3$ und $\beta 4$ in Gly m 4 durch eine Typ II- β -Schleife verbunden, während Bet v 1 an entsprechender Stelle eine größere flexible Schleife aufweist. Zweitens liegt die C-terminale α -Helix in Gly m 4 in leicht gebogener Form vor, während sie in Bet v 1 gerade verläuft. Der dritte Unterschied besteht in der Verschiebung der C-terminalen Helix entlang ihrer Längsachse relativ zum β -Faltblatt (Einzelarbeit A, Abb. 4A). In diesen drei strukturellen Merkmalen ähnelt Gly m 4 drei Bet v 1-homologen PR10-Proteinen aus der gelben Lupine, die somit interessante Kandidaten für eine Untersuchung auf Kreuzreaktivität mit Gly m 4 und Bet v 1 darstellen (Einzelarbeit A, Abb. 4B).

Durch Kombination der strukturellen Informationen mit von unseren Kooperationspartnern ermittelten immunologischen Daten bezüglich der Kreuzreaktivität von Gly m 4 mit Bet v 1 und verschiedenen homologen Allergenen konnten vier konservierte Bereiche auf der Oberfläche von Gly m 4 identifiziert werden, die potentielle kreuzreaktive IgE-Epitope darstellen (Einzelarbeit A, Abb. 6). Drei dieser Bereiche überlappen in weiten Teilen mit potentiellen kreuzreaktiven IgE-Epitopen der Allergene, die für Kreuzreaktionen auf Baumpollen der Ordnung *Fagales* verantwortlich sind. Alle vier Bereiche enthalten Aminosäuren, für die bereits durch Mimotop- oder Mutationsstudien an Bet v 1 und anderen kreuzreaktiven Homologen gezeigt wurde, dass sie die IgE-Reaktivität beeinflussen. Bei Mimotopen handelt es sich um Peptide, die nicht aufgrund ihrer Sequenz, sondern aufgrund

der Eigenschaften ihrer Aminosäuren bestimmte Oberflächenbereiche eines Proteins repräsentieren. Die Peptide, die ein konformationelles Epitop eines Allergens repräsentieren, werden aus einer Zufallsbibliothek durch Bindung der entsprechenden IgEs selektiert. Anschließend wird per computergestützter Suche der entsprechende Bereich auf der Oberfläche des Allergens lokalisiert. Die immunologischen Daten zeigten auch, dass bei gleichem strukturellen Grundgerüst eine grobe Korrelation zwischen Sequenzähnlichkeit und Grad der Kreuzreaktivität besteht, wobei Bet v 1 hier aufgrund seiner Rolle als sensibilisierendes Allergen eine Ausnahme darstellt. Kein ersichtlicher Zusammenhang besteht hingegen zwischen der thermischen Stabilität des Allergens und der Schwere der ausgelösten allergischen Reaktion, wie durch CD-Messungen an Pru av 1 und Gly m 4, das im Gegensatz zu ersterem auch schwere systemische Reaktionen hervorruft (Kleine-Tebbe *et al.*, 2002), gezeigt werden konnte. Die gewonnenen Erkenntnisse bezüglich der Lokalisation der IgE-Epitope können letztendlich zur Konstruktion hypoallergener Proteinvarianten beitragen, die in der Immuntherapie eingesetzt werden (Purohit *et al.*, 2008).

3.2 Norcoclaurin-Synthase – Ein neues Mitglied der Bet v 1-Superfamilie

Das Enzym NCS aus der gelben Wiesenraute *Thalictrum flavum* weist eine im Vergleich zu den klassischen Bet v 1-Allergenen eher geringe Sequenzidentität von 25 % mit Bet v 1 auf, und umfasst als 23 kDa Protein sowohl N- als auch C-terminale Bereiche, die keine Entsprechung in der Bet v 1-Sequenz finden. Gerade diese Unterschiede machten das Enzym zu einem umso interessanteren Kandidaten für eine strukturelle Charakterisierung.

3.2.1 Produktion des rekombinanten Proteins

Voraussetzung für Strukturuntersuchungen eines Proteins der Größe der Tf-NCS mittels NMR ist die Produktion von großen Mengen an isotoopenmarkiertem, reinem Protein. Das von Samanani *et al.* (Samanani *et al.*, 2004) konstruierte Expressionssystem *Escherichia coli* (*E.coli*) ER2566 pLysS/pET29b-*A19NCS*, das die Produktion eines die Aminosäuren 20-210 der Tf-NCS umfassenden Proteins mit N-terminalem S-Anhang und C-terminalem Hexahistidin-Anhang ermöglicht, war der Ausgangspunkt für die Optimierung der Expression, auch im Hinblick auf die notwendige Isotoopenmarkierung, und die anschließende Etablierung der Reinigung. Drei Schwachpunkte dieses Systems mussten für die Herstellung von zur Strukturbestimmung geeignetem, homogenem Protein überwunden werden. Zum

ersten verhinderten niedrige Expressionsraten die Produktion von größeren Mengen an Protein (Einzelarbeit B, Abb. 1A), zweitens erforderte die Abspaltung des bei der Strukturbestimmung hinderlichen S-Anhangs mittels Thrombin einen zusätzlichen Arbeitsschritt, und drittens erhielt man nach Thrombinspaltung und Reinigung per Nickelionen-Affinitätschromatographie zwei verschiedene *Tf*-NCS-Spezies, bei denen es sich zum einen um das korrekte die $\Delta 19$ NCS enthaltende Spaltprodukt und zum anderen um eine um weitere zehn N-terminale Aminosäuren verkürzte Version handelte (Einzelarbeit B, Abb. 3). Die Expressionsraten konnten durch die Verwendung des Bakterienstammes *E. coli* Rosetta(DE3), der für die Expression eukaryontischer Gene in *E. coli* optimiert ist, beträchtlich erhöht werden (Einzelarbeit B, Abb. 1C). Die Thrombinspaltung und das Auftreten zweier *Tf*-NCS-Spezies konnte durch die Konstruktion eines Expressionsvektors, der nur für die Aminosäuren 30-210 der *Tf*-NCS mit C-terminalem Hexahistidin-Anhang kodiert, umgangen werden. Mit dem neuen Expressionssystem *E. coli* Rosetta(DE3)/ pET29b- $\Delta 29$ NCS konnten auch nach Anzucht in Minimalmedium, die im Zuge der Isotopenmarkierung nötig ist, hohe Ausbeuten an reinem, strukturell homogenem Protein erzielt werden (Einzelarbeit B, Abb. 4 und 5). Auch die enzymatische Aktivität des verkürzten Proteins $\Delta 29$ NCS sowie das Vorhandensein der Tertiärstruktur konnte bestätigt werden (Einzelarbeit B, Abb. 6 und 7).

Parallel zur Entwicklung dieses Expressionssystems wurde zum Zwecke enzymkinetischer und mechanistischer Untersuchungen auch von Luk *et al.* (Luk *et al.*, 2007) eine Optimierung des von Samanani *et al.* (Samanani *et al.*, 2004) publizierten Expressionssystems vorgenommen. Zur Erhöhung der Expressionsrate wurde hier mit der Wahl des *E. coli*-Stammes BL21(DE3)RIL die gleiche Strategie angewendet, da auch dieser Stamm für die Expression eukaryontischer Gene in *E. coli* optimiert ist. Es wurde jedoch in diesem Fall weder eine Umklonierung vorgenommen, noch wurde der S-Anhang mittels Thrombinspaltung entfernt. Interessanterweise tritt vermutlich auch bei diesem Expressionssystem nach der Reinigung eine zweite Spezies auf, die sich bei der Analyse des gereinigten Proteins mit Natriumdodecylsulfat-Polyacrylamid-Gelelektrophorese (SDS-PAGE) in einer schwachen Bande unterhalb der $\Delta 19$ NCS-Bande manifestiert (Luk *et al.*, 2007).

Erst kürzlich wurde noch ein weiteres alternatives Expressions- und Reinigungsprotokoll mit dem Ziel der strukturellen Charakterisierung der *Tf*-NCS mittels Röntgenstrukturanalyse publiziert (Pasquo *et al.*, 2008). Hier bestand die Strategie zur Erhöhung der Expressionsrate in der Konstruktion eines synthetischen Gens für $\Delta 19$ NCS, dessen Sequenz an die *codon*

usage von *E. coli* angepasst wurde. Die Analyse des gereinigten Proteins per SDS-PAGE zeigt, dass mit dem verwendeten Expressionssystem *E. coli* BL21(DE3)/pET22b-*Δ19NCS* und dem angewendeten Reinigungsprotokoll nur eine Proteinspezies erhalten wird.

Zum Zwecke der Strukturbestimmung und zur Durchführung quantitativer Bindungsstudien mittels NMR war das von uns etablierte Protokoll am besten geeignet, da es die Produktion des mit 190 Aminosäuren kleinsten, aber dennoch enzymatisch aktiven Proteinkonstruktes in isotoopenmarkierter Form ermöglicht.

3.2.2 Der Oligomerisierungszustand und die methodischen Konsequenzen

Der Oligomerisierungszustand eines Enzyms kann eine wichtige Rolle für dessen Funktionalität und Regulation spielen. Da der Oligomerisierungszustand der *Tf*-NCS bisher nur für das aus der Pflanze isolierte Enzym ermittelt wurde, und dies zu einem Zeitpunkt geschah, als die Aminosäuresequenz und somit das tatsächliche theoretische Molekulargewicht noch unbekannt war (Samanani & Facchini, 2001), wurde der Oligomerisierungszustand der rekombinanten *Tf*-NCS in dieser Arbeit mit Hilfe von zwei unabhängigen Methoden untersucht. Sowohl die Größenausschlusschromatographie als auch die NMR-spektroskopischen Relaxationsmessungen zeigen eine Konzentrationsabhängigkeit des Oligomerisierungszustandes (Einzelarbeit C, Tab. 1). Bei niedrigen Enzymkonzentrationen, wie sie für kinetische Messungen eingesetzt wurden (Luk *et al.*, 2007), liegt ein Großteil des Proteins als Monomer vor. Die Zugabe von Substrat hat keinen Einfluss auf den Oligomerisierungszustand. Diese Beobachtungen widersprechen der Vermutung, dass die bezüglich Dopamin gemessene Kooperativität der Kinetik durch eine Dimerisierung des Enzyms zustande kommt (Luk *et al.*, 2007; Samanani *et al.*, 2004).

Die Konzentrationsabhängigkeit des Oligomerisierungszustandes beeinflusste auch die weitere methodische Vorgehensweise. Da es sich bei der NMR-Spektroskopie um eine relativ insensitive Methode handelt, werden zur Strukturbestimmung Proben mit hoher Proteinkonzentration benötigt. Zudem limitiert die Größe des Proteins die Auswertbarkeit der Spektren, da die proportional zur molekularen Masse ansteigende transversale Relaxation zu steigendem Signalverlust führt (Clare & Gronenborn, 1998). Im Fall der *Tf*-NCS bringt die steigende Oligomerisierungstendenz bei steigender Konzentration mit sich, dass sich das Protein bei einer Konzentration von 1 mM bezüglich der Relaxation nicht wie ein 21 kDa-Protein, sondern wie ein 29 kDa-Protein verhält (Einzelarbeit C, Tab. 1). Deshalb wurden für die Zuordnung der Amidprotonenresonanzen des Proteinerückgrats zwei Techniken

angewendet, die häufig bei Systemen größer 25 kDa genutzt werden, um dem Signalverlust durch Relaxationseffekte entgegenzuwirken (Clore & Gronenborn, 1998). Zum einen wurde die *Tf*-NCS perdeutert, d.h. durch Expression des Proteins in 100 % D₂O und anschließender Lösung des lyophilisierten Proteins in 90 % H₂O/10 % D₂O wurden alle nicht mit dem Lösungsmittel austauschenden Protonen durch Deuteronen ersetzt. Aufgrund ihres im Vergleich zum Proton geringeren gyromagnetischen Verhältnisses reduzieren die Deuteronen die über Dipol-Dipol-Wechselwirkungen vermittelte transversale Relaxation und bewirken somit einen Gewinn an Signalintensität (Gardner & Kay, 1998). Zum anderen wurde für die Messung der dreidimensionalen (3D-) Zuordnungsexperimente die TROSY-Technik (*Transverse Relaxation Optimized Spectroscopy*) angewendet, deren Pulssequenzen für die Minimierung von Signalverlust durch transversale Relaxation optimiert sind (Pervushin *et al.*, 1997; Fernandez & Wider, 2003).

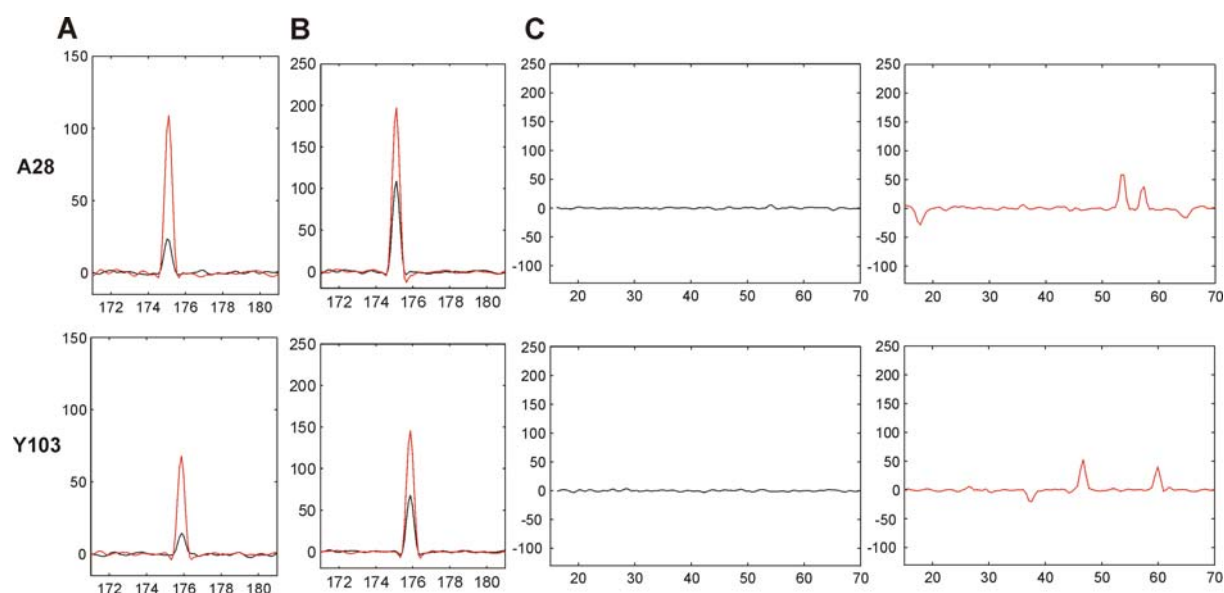


Abb. 9: Einfluss von Deuterierung und TROSY auf die Sensitivität verschiedener 3D-NMR-Experimente

Die beiden Aminosäuren A28 und Y103 wurden zufällig als Beispiele gewählt; die Messtemperatur bei allen Messungen betrug 305,5 K; die Proteinkonzentration aller Proben betrug 400 µM.

A) Signalintensität im HNCO-Spektrum in Abhängigkeit von der Deuterierung: Schwarz = ¹³C, ¹⁵N-markierte *Tf*-NCS-Probe (Feldstärke: 700 MHz, Messzeit: 24 h); Rot = Perdeuterte ¹³C, ¹⁵N-markierte *Tf*-NCS-Probe (Feldstärke: 800 MHz, Messzeit: 27 h)

B) Signalintensität bei Verwendung einer perdeuterten ¹³C, ¹⁵N-markierten *Tf*-NCS-Probe in Abhängigkeit von der Messtechnik: Schwarz = konventionelles HNCO (Feldstärke: 800 MHz, Messzeit: 27 h); Rot = TROSY-HNCO (Feldstärke: 800 MHz, Messzeit: 27 h)

C) Signalintensität im HNCACB-Spektrum in Abhängigkeit von Deuterierung und Messtechnik: Schwarz = konventionelles HNCACB-Spektrum einer ¹³C, ¹⁵N-markierten *Tf*-NCS-Probe (Feldstärke: 700 MHz, Messzeit: ca. 4,5 d); Rot = TROSY-HNCACB-Spektrum einer perdeuterten ¹³C, ¹⁵N-markierten *Tf*-NCS-Probe (Feldstärke: 800 MHz, Messzeit: 4,5 d)

Allein durch die Verwendung perdeutrierter Proben konnten die Signalintensitäten im 3D-HNCO ungefähr vervierfacht werden (Abb. 9A). Während die Anwendung der TROSY-Technik auf nicht-deuterte Proben kaum zur Verbesserung der Sensitivität beitrug, konnte durch Kombination mit der Deuterierung im TROSY-HNCO ein weiterer Signalintensitäts-

gewinn im Rahmen des für diese Technik erwarteten Faktors von ca. 2,4 erzielt werden (Salzmann *et al.*, 1998; Abb. 9B). Im sehr insensitiven 3D-HNCACB-Experiment ermöglicht die Kombination dieser beiden Techniken erst die Interpretation der Spektren, da die Signale vorher kaum vom Rauschen zu unterscheiden sind (Abb. 9C). Letztendlich konnte dadurch die Rückgrat-Zuordnung von 86 % der Aminosäuren erzielt werden (Einzelarbeit C, Abb. 4). Der nächste Schritt auf dem Weg zur dreidimensionalen Struktur ist die Zuordnung der Seitenketten. Da in einem perdeutierten Protein die an Kohlenstoff gebundenen Protonen der Seitenketten durch Deuteronen ersetzt sind, die in den entsprechenden Korrelationsexperimenten kein Signal liefern, wurden die Messungen mit nur partiell deutiertem Protein durchgeführt, das in 50 % D₂O exprimiert wurde. Weder dieser Kompromiss aus einer begrenzten Anzahl an signalliefernden Protonen, die jedoch gleichzeitig zur Verringerung der transversalen Relaxationsrate beiträgt, noch die Verwendung von nicht-deutiertem Protein für die Messungen führte jedoch zu Spektren mit ausreichender Qualität für eine Seitenkettenzuordnung.

3.2.3 Das semi-experimentelle Homologiemodell

Eine ausschließlich auf experimentellen Daten basierende Bestimmung der Struktur der *Tf*-NCS war aufgrund der fehlenden Seitenkettenzuordnung nicht möglich. Deshalb wurde das Strukturmodell basierend auf einer Kombination der vorliegenden experimentellen Daten mit der Technik der Homologiemodellierung erstellt. Während der Vergleich des mittels CD-Spektroskopie geschätzten Sekundärstrukturgehalts und des Schmelzpunktes die Wahl von Bet v 1 als Modellierungsvorlage unterstützte (Einzelarbeit C, Abb. 3), gingen aus dem *chemical shift index* (CSI) gewonnene Informationen bezüglich der Lokalisation der Sekundärstrukturelemente und aus NOESY-Experimenten abgeleitete Informationen zur relativen Orientierung der enthaltenen Faltblattstränge in die Konstruktion und Validierung des Strukturmodells mit ein (Einzelarbeit C, Abb. 5A). Anhand dieses Strukturmodells konnte erstmals gezeigt werden, dass auch die *Tf*-NCS die typische Bet v 1-Faltungstopologie aufweist und somit zur Bet v 1-Superfamilie zählt (Einzelarbeit C, Abb. 5B). Der einzige Unterschied findet sich im Bereich der C-terminalen α -Helix, die im Gegensatz zu Bet v 1 in der *Tf*-NCS von zwei weiteren kurzen α -Helices umgeben ist.

3.2.4 Vergleich des Homologiemodells mit der Kristallstruktur

Erst kürzlich wurde unser semi-experimentelles Homologiemodell durch die Lösung der Kristallstruktur der Tf-NCS (Ilari *et al.*, 2009) in weiten Teilen bestätigt (Abb. 10). Vor allem im Bereich des siebensträngigen antiparallelen β -Faltblatts sind keine wesentlichen Unterschiede in der Konformation des Proteinrückgrates zu erkennen, was sich in einem relativ geringen Wert der Standardabweichung von 1,4 Å für die Überlagerung der β -Stränge äußert. Selbst einige Schleifen in diesem Bereich, wie z.B. die Verbindung zwischen Strang $\beta 4$ und $\beta 5$ oder die $\beta 2$ und $\beta 3$ verknüpfende P-Schleife, weisen im Modell die gleiche oder eine sehr ähnliche Konformation wie in der Kristallstruktur auf. Auch die beiden V-förmig angeordneten Helices $\alpha 1$ und $\alpha 2$ sind in der Kristallstruktur in ähnlicher Orientierung vorhanden, ebenso wie der kurze helikale Bereich $\alpha 3$.

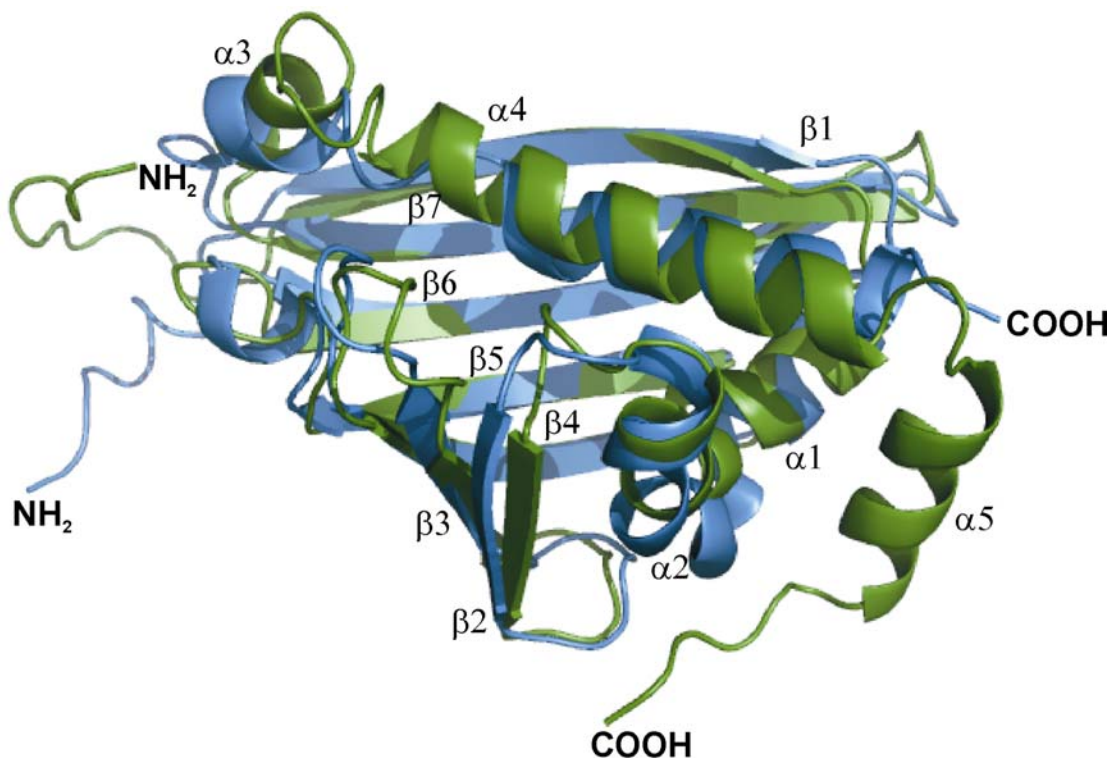


Abb. 10: Vergleich des semi-experimentellen $\Delta 29$ NCS-Modells (grün) mit der $\Delta 19$ NCS-Kristallstruktur (blau; Ilari *et al.*, 2009, PDB-Code 2VNE)

Der C-terminal von $\alpha 3$ gelegene Bereich unterscheidet sich im Modell und in der Kristallstruktur. Während der direkt an $\alpha 3$ anschließende Bereich zwischen P145 und P151 im Modell eine größere Schleife ausbildet, verläuft das Proteinrückgrat in der Kristallstruktur parallel zur Längsachse der folgenden Helix $\alpha 4$. Die beiden helikalen Bereiche sind folglich in der Kristallstruktur weniger als zwei getrennte Helices denn als eine unterbrochene lange Helix zu betrachten, die der langen C-terminalen Helix von Bet v 1 entspricht. Dies führt dazu, dass die Sequenz in den wiederum in ähnlicher Orientierung vorliegenden Helices $\alpha 4$

um acht Aminosäuren, d.h. ca. zwei Helixwindungen verschoben ist. Die Differenzen zur Kristallstruktur in diesem Bereich des Modells sind zum einen auf die im Vergleich zu anderen Teilen des Proteins geringere Sequenzähnlichkeit mit Bet v 1 zurückzuführen, die größere Freiheit für das Sequenzalignment zulässt, das dem Modell zugrunde liegt. Zum anderen konnten aufgrund des für die Bet v 1-Faltung typischen Hohlraums zwischen der langen C-terminalen α -Helix und dem β -Faltblatt und der fehlenden Seitenketten-Zuordnung keine experimentellen Daten zur relativen Orientierung dieser beiden Sekundärstrukturelemente gewonnen werden.

Der Sequenzbereich, in dem im Modell Helix $\alpha 5$ zu finden ist, konnte vermutlich aufgrund seiner Flexibilität in der Kristallstruktur nicht aufgelöst werden. Auch die N-terminalen Aminosäuren sind in der Kristallstruktur nur in Teilen sichtbar. Sowohl im Modell als auch in der Kristallstruktur geht der jeweils sichtbare Bereich N-terminal des Stranges $\beta 1$ kaum Wechselwirkungen mit dem kompakt gefalteten Teil des Proteins ein. Auch die im $[^1\text{H}, ^{15}\text{N}]$ -TROSY-Spektrum fehlenden Resonanzen der Aminosäuren am N-Terminus der $\Delta 29\text{NCS}$ sprechen für hohe Flexibilität dieses Bereichs.

3.2.5 Substratbindung unter Berücksichtigung der Komplexstruktur

Zur Identifizierung der Substratbindungsstellen und somit des aktiven Zentrums des Enzyms mittels NMR wurde die Methode der *chemical shift perturbation* (CSP) angewendet. Hierbei wird die Veränderung der chemischen Umgebung beteiligter Aminosäuren durch Ligandenbindung genutzt, die sich im $[^1\text{H}, ^{15}\text{N}]$ -HSQC-Spektrum in der Verschiebung der entsprechenden Resonanz äußert, sofern der Austausch zwischen freiem und gebundenem Zustand bezogen auf die NMR-Zeitskala schnell stattfindet (Zuiderweg, 2002). Die Verschiebung einer Resonanz kann somit entweder von direktem Kontakt der entsprechenden Aminosäure bzw. deren Rückgrat-Amidgruppe mit dem Liganden oder von indirekten Effekten wie einer liganden-induzierten Konformationsänderung, der Reorientierung von Seitenketten oder von Ringstromeffekten aromatischer Liganden verursacht werden (McCoy & Wyss, 2000).

Da neben der Kristallstruktur des freien Zustands der *Tf*-NCS auch die Komplexstruktur mit dem Substrat Dopamin und dem nicht-reaktiven Substratanalogon 4-Hydroxybenzaldehyd (HBA) gelöst werden konnte (Ilari *et al.*, 2009), werden im Folgenden die Ergebnisse der CSP auch im Hinblick auf diese Komplexstruktur diskutiert. Weil die Nummerierung der Aminosäuren im Modell auf das verwendete Konstrukt bezogen ist, in der Kristallstruktur

jedoch auf das Volllängenprotein, ist zusätzlich zur Nummerierung des Modells in Klammern die Nummerierung der Kristallstruktur angegeben. Während durch Zugabe des 4-HPAA-Analogons 4-Hydroxyphenylessigsäure-Methylester in weiten Teilen des Proteins Änderungen der chemischen Verschiebung induziert werden, beschränken sich die signifikanten Änderungen im Fall von Dopamin hauptsächlich auf die zwei Bereiche um F71(99) und um M155(183) (Einzelarbeit C, Abb. 6 A-D).

Ein Vergleich des Proteinrückgrates der Kristallstruktur ohne Substrate mit dem Rückgrat der Komplexstruktur mit dem 4-HPAA-Analogon HBA allein bzw. mit HBA und Dopamin ergibt eine sehr geringe Standardabweichung der Atomkoordinaten von 0,3-0,5 Å. Somit wird deutlich, dass die Substratbindung zu keiner nennenswerten Änderung der globalen Struktur führt. Dies unterstützt unsere Vermutung, dass die Änderungen der chemischen Verschiebung in weiten Teilen des Proteins bei Zugabe des 4-HPAA-Analogons 4-Hydroxyphenylessigsäure-Methylester im Wesentlichen durch die Umorientierung von Seitenketten hervorgerufen wird. Die Betrachtung der Seitenketten bestätigt diese Vermutung zumindest teilweise. Im Komplex mit Dopamin und HBA liegt im Fall von F84(112) ein anderes Rotamer vor als in der freien Struktur (Abb. 11).

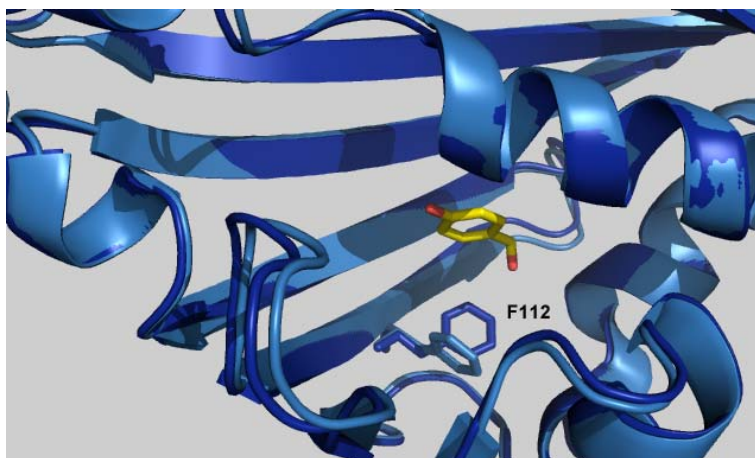


Abb. 11: Umorientierung der Seitenkette von F112 bei Substratbindung

Hellblau = *Tf*-NCS ohne Substrate (PDB-Code 2VNE); Dunkelblau = *Tf*-NCS im Komplex mit Dopamin und HBA (davon nur HBA in gelb dargestellt, PDB-Code 2VQ5); Nummerierung laut Kristallstruktur

Da durch den Ringstromeffekt des aromatischen Systems besonders weitreichende Änderungen der chemischen Umgebung verursacht werden (McCoy & Wyss, 2000), sind vermutlich sowohl die Änderungen der chemischen Verschiebung in der unmittelbaren Umgebung von F84(112), also im Faltblattstrang β_4 , als auch jene im Bereich des benachbarten β -Stranges 5 und der in räumlicher Nähe gelegenen Helix α_2 , die das deutlich beeinflusste W37(65) enthält, auf diese Umorientierung zurückzuführen. Bemerkenswert ist, dass F84(112) in dem Molekül der Komplexstruktur, das nur HBA gebunden hat, in der gleichen Konformation vorliegt wie im freien Zustand. Dies deutet darauf hin, dass F84(112)

erst bei Bindung beider Substrate bzw. bei Bindung von Dopamin die Orientierung ändert. Dagegen sprechen jedoch die unter der Signifikanzgrenze liegenden Änderungen der chemischen Verschiebung in den entsprechenden Bereichen bei Zugabe von Dopamin sowohl in An- als auch in Abwesenheit von 4-HPAA. Möglicherweise ist dieser Widerspruch auf die unterschiedlichen Bedingungen während der Kristallisation im Vergleich zur NMR-Titration zurückzuführen. Die aus Produktinhibitionskinetiken gewonnene Schlussfolgerung, dass 4-HPAA vor Dopamin bindet (Samanani & Facchini, 2002), unterstützt die These, dass 4-HPAA die Umorientierung induziert.

Die durch Dopamin sowie die durch das 4-HPAA-Analogon induzierten Änderungen um F71(99) und um M155(183) hingegen sind hauptsächlich auf direkten Kontakt der entsprechenden Aminosäuren mit dem jeweiligen Liganden zurückzuführen. Aufgrund der in Abschnitt 3.2.4. erwähnten Verschiebung um ca. 2 Windungen im Bereich der Helix $\alpha 4$ rückt M155(183) vom N-Terminus der Helix im Modell in den mittleren Teil der Helix in der Kristallstruktur. Die Position und die Ausrichtung von F71(99) im Modell hingegen entspricht nahezu exakt der Kristallstruktur (Abb. 12A). Die beiden Liganden binden unter gegenseitiger Stabilisierung durch Stapelung der nahezu parallel gelagerten aromatischen Ringsysteme in den hauptsächlich hydrophoben Hohlraum im Inneren der Tf-NCS, wobei sie sowohl hydrophile als hydrophobe Wechselwirkungen mit mehreren Aminosäuren eingehen.

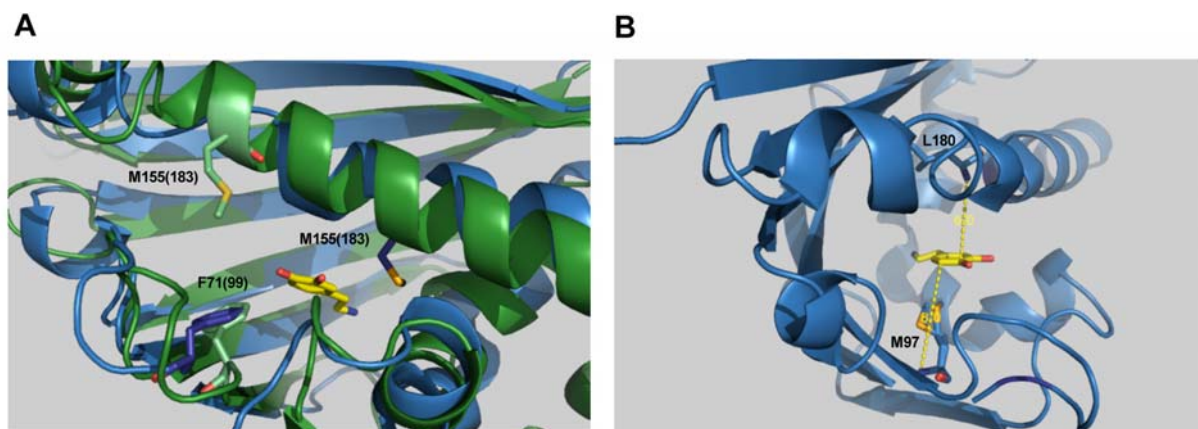


Abb. 12: Bindung von Dopamin

A) Position der aufgrund des Modells und der CSP ermittelten an der Dopaminbindung beteiligten Reste F71(99) und M155(183) im Tf-NCS-Modell (grün) und in der Komplexstruktur (PDB-Code 2VQ5, blau) relativ zum gebundenen Dopamin (gelb)

B) Abstände der Rückgratamidgruppen der beiden Aminosäuren M69(97) und L152(180), die eine sehr hohe CSP aufweisen, vom in der Komplexstruktur (PDB-Code 2VQ5) gebundenen Dopamin (gelb); Abstände in Å; Nummerierung laut Kristallstruktur

Im Bereich um F71(99) interagiert nicht das Phenylalanin selber, sondern das benachbarte M69(97) über hydrophobe Wechselwirkungen mit Dopamin. Da in der Komplexstruktur zwei alternative Rotamere für M69(97) angegeben sind, besteht auch hier die Möglichkeit einer ligandeninduzierten Umorientierung. Die extrem starken Änderungen der chemischen

Verschiebung von M69(97) und dem im Bereich um M155(183) ebenfalls auf der substratzugewandten Seite der Helix $\alpha 4$ lokalisierten L152(180) rühren vermutlich daher, dass diese Aminosäuren und insbesondere ihre für das Messsignal verantwortlichen Rückgrat-Amidgruppen auf einer Achse liegen, die das Ringsystem von Dopamin senkrecht in der Mitte schneidet (Abb. 12B). Sowohl M155(183) als auch das zu L152(180) benachbarte P151(179), dessen CSP aufgrund des fehlenden Rückgrat-Amidprotons nicht detektierbar ist, tragen durch hydrophobe Wechselwirkungen zur Bindung von Dopamin bei. Der einzige hydrophile Beitrag zur Bindung von Dopamin besteht in der Ausbildung einer Wasserstoffbrückenbindung zwischen der C1-Hydroxylgruppe und der Hydroxylgruppe von Y80(108). Dieses Tyrosin zeigt jedoch keine signifikante CSP, da das Rückgrat-Amidproton weit von Dopamin entfernt ist. Ein weiterer wichtiger Rest, der hydrophob-polare Wechselwirkungen mit Dopamin eingeht, ist E82(110).

Auch das 4-HPAA-Analogon HBA interagiert über hydrophob-polare Wechselwirkungen mit den beiden Methioninen M69(97) und M155(183), sowie dem benachbarten L152(180), was sich im CSP mit signifikanten Änderungen für alle drei Aminosäuren äußert. Die extrem hohe CSP von M155(183) wird vermutlich durch Beiträge seiner Seitenkettenumorientierung sowie der Positionierung relativ zum aromatischen Ring des 4-HPAA-Analogons verursacht. Während im Molekül, welches allein das 4-HPAA-Analogon gebunden hat, Wasserstoffbrücken zwischen der Carbonyl- bzw. der Hydroxylgruppe von HBA und K94(122) bzw. E82(110) ausgebildet werden, interagiert die Hydroxylgruppe von HBA im Komplex mit beiden Substraten mit D113(141). Außer für E82(110) ist auch für diese Aminosäuren eine signifikante CSP zu beobachten. Weiterhin interagiert das 4-HPAA-Analogon in Abwesenheit von Dopamin mit dem in Helix $\alpha 2$ lokalisierten L40(68), welches während der Titration intermediäres Austauschverhalten zeigt. Auch das bereits bezüglich der Seitenkettenumorientierung erwähnte F84(112) zeigt hydrophobe und hydrophob-polare Wechselwirkungen mit dem 4-HPAA-Analogon.

3.2.6 Katalytischer Mechanismus

Neben der NCS katalysiert auch die Strictosidin-Synthase eine ähnliche Pictet-Spengler-Reaktion, in diesem Falle die Kondensation des Aldehyds Secologanin und des Amins Tryptamin zu Strictosidin, dem Vorläufer aller pflanzlichen Indol-Alkaloide (Kutchan, 1993). Die Kristallstruktur dieses Enzyms im Komplex mit den Substraten zeigt, dass während der Katalyse die Carboxylgruppe eines Glutamates eine wichtige Rolle spielt (Ma *et al.*, 2006).

Aufgrund der Analogie der katalysierten Reaktion wurde E75(103), dessen Seitenkette in unserem NCS-Modell zwischen den bereits im letzten Absatz aufgrund ihrer signifikanten CSP diskutierten Aminosäuren F71(99) und M155(183) liegt, als katalytischer Rest der NCS vorgeschlagen (Einzelarbeit C, Abb. 6 E). Durch die Verschiebung der Aminosäuresequenz des Modells im Bereich der Helix $\alpha 4$ im Vergleich zur Kristallstruktur vergrößert sich jedoch die räumliche Entfernung zwischen F71(99) und M155(183). Bezieht man weiterhin die Lage der Substrate bzw. Substratanaloga in der Komplexstruktur mit ein, so ist eine Beteiligung von E75(103) an der Katalyse der Reaktion unwahrscheinlich.

Die Analyse der Komplexstruktur bezüglich der Position bestimmter Reste relativ zu den Substraten kombiniert mit Messungen der enzymatischen Aktivität dreier Mutanten, in denen die potentiell katalytisch aktiven Aminosäuren Y80(108), E82(110) und K94(122) ausgetauscht wurden (Abb. 13A), deutet auf den in Abb. 13B schematisch dargestellten Reaktionsmechanismus hin (Ilari *et al.*, 2009).

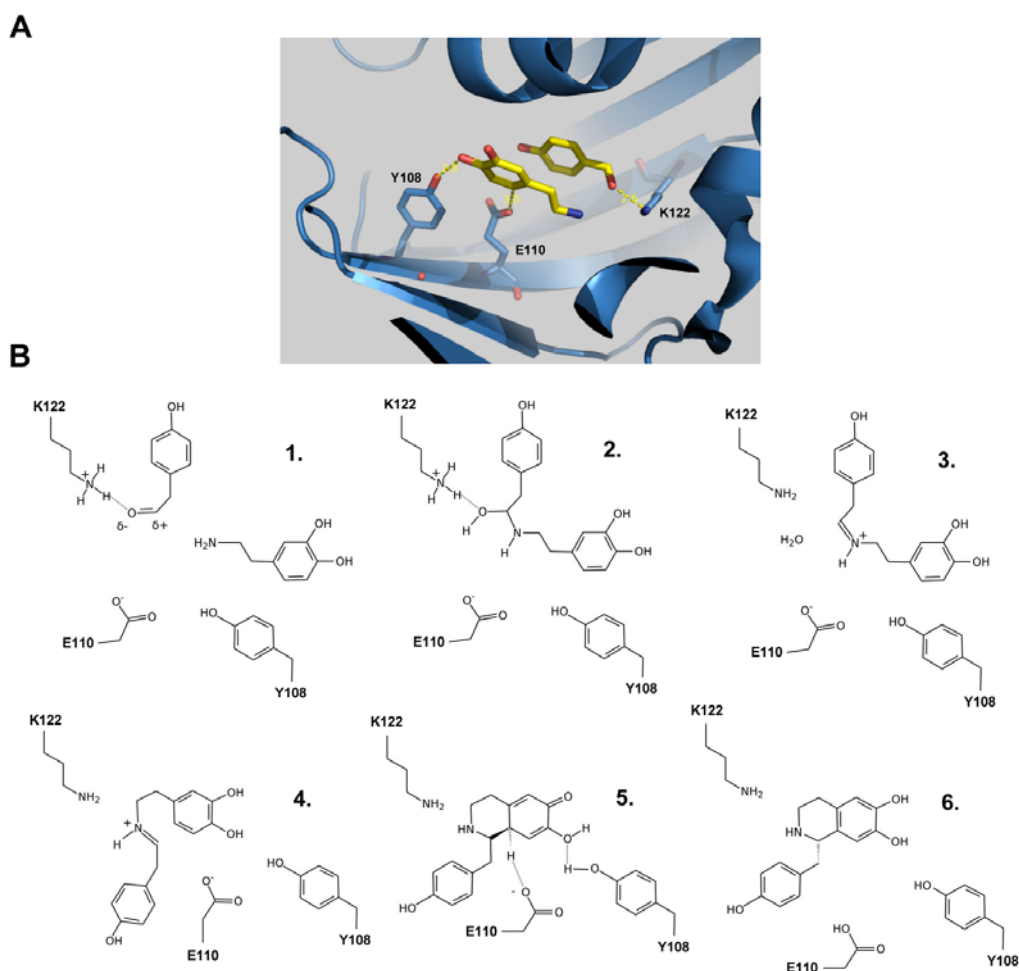


Abb. 13: Potentielle katalytische Reste und Reaktionsmechanismus der *Tf*-NCS

A) Relative Lage der potentiellen katalytischen Reste der *Tf*-NCS zu den in der Komplexstruktur gebundenen Substraten bzw. Substratanaloga Dopamin und HBA; Abstände in Å; Nummerierung laut Kristallstruktur (PDB-Code 2VQ5)

B) Aufgrund der Komplexstruktur vorgeschlagener Reaktionsmechanismus der *Tf*-NCS (nach Ilari *et al.*, 2008)

Dabei wird die Reaktion im Gegensatz zum Mechanismus der Strictosidin-Synthase nicht von einer Deprotonierung der Aminogruppe des Dopamins durch einen Glutamatrest eingeleitet, sondern die Bildung des Iminiumions wird durch die Protonierung der Carbonylgruppe von 4-HPAA erleichtert (Abb. 13B, Schritt 1 und 2). Als Protonendonator fungiert hier die Aminogruppe des K94(122), die vermutlich auch an der zur Bildung des Iminiumions nötigen Abspaltung eines Wassermoleküls beteiligt ist (Abb. 13B, Schritt 3). Der folgende Schritt erfordert eine Reorientierung des Iminiumions (Abb. 13B, Schritt 4), gefolgt von einer elektrophilen Substitution am C5-Atom des Dopamins, die zum Ringschluss führt (Abb. 13B, Schritt 5). Wie bereits von Luk *et al.* (2007) vorgeschlagen (s. Abschnitt 1.1.4.), wird dieser Schritt durch die Ausbildung eines Phenolations an der C2-Hydroxylgruppe des Dopamins begünstigt, deren Proton vermutlich von einem Wassermolekül akzeptiert wird. Weiterhin könnte die Wasserstoffbrückenbindung zwischen Y80(108) und der C1-Hydroxylgruppe des Dopamins zur Stabilisierung des entstandenen σ -Intermediats beitragen. Die anschließende Wiederherstellung des aromatischen Zustandes durch Abstraktion des C5-Protons der Katecholeinheit wird mit hoher Wahrscheinlichkeit durch die Carboxylgruppe von E82(110) unterstützt (Abb. 13B, Schritt 6). Die Position dieses Restes in Kombination mit der sterischen Restriktion durch das Enzym während der Umorientierung des Iminiumions garantiert für die Stereospezifität der Reaktion.

Während die Komplexstruktur also den von Luk *et al.* (2007) vorgeschlagenen Reaktionsmechanismus weitgehend bestätigt, kann sie jedoch keine Erklärung für die beobachtete Kooperativität der Enzymkinetik bezüglich Dopamin liefern (Luk *et al.*, 2007; Samanani *et al.*, 2004; Samanani & Facchini, 2002). Ein Vergleich der Kristallstruktur ohne Substrate mit der Komplexstruktur zeigt, dass sich ligandeninduzierte Strukturänderungen intramolekular auf die Umorientierung weniger Seitenketten beschränken. Bezüglich der Quartärstruktur bestätigt der Vergleich das Ergebnis unserer zum Oligomerisierungszustand der Tf-NCS durchgeführten Untersuchungen (s. Abschnitt 3.2.2.), da auch in der Kristallstruktur keine durch Ligandenbindung induzierte Oligomerisierung beobachtet wird.

3.2.7 NCS-Aktivität und die physiologische Funktion der Bet v 1-Familie

Die große Ähnlichkeit der dreidimensionalen Struktur der NCS mit der Struktur „klassischer“ Mitglieder der Bet v 1-Familie bzw. der PR10-Proteine und vor allem die Konservierung des zur Ligandenbindung genutzten hydrophoben Hohlraumes wirft die Frage auf, ob die bis dato ungeklärte physiologische Funktion dieser Proteine ebenfalls in der Katalyse der NCS-

Reaktion besteht. Neben unserem $\Delta 29\text{NCS}$ -Konstrukt (Einzelarbeit A, Abb. 6) wurden deshalb auch rekombinantes Bet v 1 und Pru av 1 auf die gleiche Weise einem Test auf NCS-Aktivität unterzogen (Facchini, unveröffentlicht). Dieser ergab jedoch, dass die beiden Allergene keine NCS-Aktivität zeigen. Auch die Titration von Bet v 1 mit Dopamin ergab keine signifikanten Änderungen der chemischen Verschiebung der Resonanzen im $[^1\text{H}, ^{15}\text{N}]$ -HSQC-Spektrum (Berkner, unveröffentlicht). Folglich ist Bet v 1 weder in der Lage, die NCS-Reaktion zu katalysieren, noch bindet es eines der Substrate. Die physiologische Funktion der PR10-Proteine bleibt also weiterhin ungeklärt. Ein erst kürzlich durchgeführter Vergleich der Funktionen aller in der PDB hinterlegten Proteine, die eine Bet v 1-ähnliche Struktur aufweisen, zeigt jedoch, dass die gemeinsame Aufgabe des konservierten hydrophoben Hohlraums in der Bindung hydrophober, meist sperriger Liganden besteht (Radauer *et al.*, 2008b).

4 Ausblick: *Epitope grafting*

Im Rahmen dieser Arbeit konnte geklärt werden, dass die *Tf*-NCS trotz ihrer geringen Sequenzähnlichkeit zu Bet v 1 eine nahezu identische Faltungstopologie besitzt. Aufgrund dieser Ähnlichkeit wurden immunologische Tests durchgeführt, die ergaben, dass das Enzym keine bis sehr geringe Kreuzreaktivität mit Bet v 1 aufweist (Vieths, unveröffentlicht). Diese Eigenschaften machen das Protein zu einem interessanten Kandidaten für das sogenannte *Epitope grafting*, eine spezielle Form des *Protein grafting* (Fernandez-Carneado *et al.*, 2000; Sia & Kim, 2003). Bei dieser Methode wird versucht, die Bindungsstelle eines Proteins für einen bestimmten Antikörper auf ein anderes Protein zu übertragen. In diesem Fall soll ein potentielles IgE-Epitop des Allergens Bet v 1 auf die *Tf*-NCS übertragen werden. Zeigt die dementsprechend mutierte *Tf*-NCS Kreuzreaktivität mit Bet v 1, so können daraus sehr detaillierte Informationen über die sequentiellen und strukturellen Voraussetzungen für die IgE-Bindung gewonnen werden.

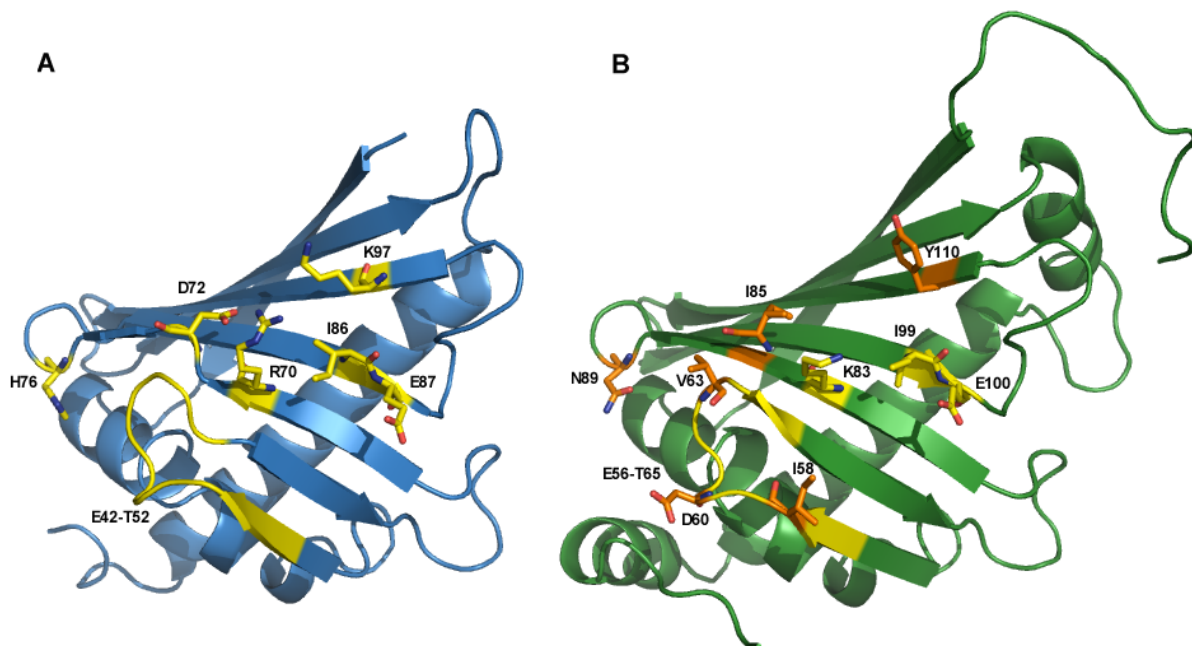


Abb 14: Übertragung eines potentiellen IgE-Epitops von Bet v 1 auf $\Delta 29$ NCS

A) Potentielles IgE-Epitop (gelb) des Hauptbirkenpollenallergens Bet v 1 (PDB-Code 1BV1, Mirza *et al.*, 2000)
 B) Entsprechende Aminosäuren der strukturhomologen $\Delta 29$ NCS: Gelb = identische oder ähnliche Aminosäure; Orange = Durch Mutation an Bet v 1 anzupassende Aminosäure; Im Bereich der P-Schleife findet sich in $\Delta 29$ NCS keine Entsprechung zu N43 von Bet v 1

Mit diesem Ziel wurden bereits während dieser Arbeit mehrere Varianten der $\Delta 29$ NCS konstruiert und auf Kreuzreaktivität getestet. Die eingeführten Mutationen basieren auf der Komplexstruktur von Bet v 1 mit dem Fab-Fragment eines IgG-Antikörpers, der die Bindung spezifischer IgE-Antikörper aus Birkenpollenallergiker-Patientenseren inhibiert (Mirza *et al.*, 2000). Das potentielle Epitop besteht hauptsächlich aus der P-Schleife und direkt

anschließenden Aminosäuren, sowie aus einzelnen Aminosäuren im β -Faltblatt (Abb. 14A). Da in diesem Bereich das $\Delta 29$ NCS-Modell vor allem bezogen auf die Aminosäurepositionen nahezu vollständig mit der NCS-Kristallstruktur übereinstimmt, eignete es sich hervorragend als Grundlage für das Design der Mutationen (Abb. 14B). Tab. 2 gibt eine Übersicht über die anhand dieses strukturbasierten Sequenzvergleichs eingeführten Mutationen.

Tab. 2: Zum Zwecke des *Epitope grafting* in $\Delta 29$ NCS eingeführte Mutationen

Bezeichnung	Eingeführte Mutationen
$\Delta 29$ NCS_1M	I58E
$\Delta 29$ NCS_3M	I58E, D60N, V63P
$\Delta 29$ NCS_4M	I58E, D60N, V63P, N57*
$\Delta 29$ NCS_5M	I58E, D60N, V63P, N57*, I85D
$\Delta 29$ NCS_6Ma	I58E, D60N, V63P, N57*, I85D, N89H
$\Delta 29$ NCS_6Mb	I58E, D60N, V63P, N57*, I85D, Y110K
$\Delta 29$ NCS_7M	I58E, D60N, V63P, N57*, I85D, N89H, Y110K

*Nach Insertion von N an Position 57 wurde für die Bezeichnung der Mutationen die ursprüngliche Nummerierung der C-terminal davon gelegenen Reste beibehalten.

Die Untersuchung aller mutierten Proteine auf eine intakte Sekundär- und Tertiärstruktur ergab, dass die im Bereich der P-Schleife eingeführten Mutationen keinen oder nur geringen Einfluss auf Sekundär- und Tertiärstruktur hatten, während die Mutationen im Bereich des β -Faltblatts zu einer mit der Zahl der Mutationen ansteigenden Aggregationstendenz führten (Berkner, unveröffentlicht; Chrobok, 2008). Für die siebenfach mutierte $\Delta 29$ NCS_7M konnte keine geordnete Tertiärstruktur mehr nachgewiesen werden. Mit Ausnahme von $\Delta 29$ NCS_7M sollten die mutierten Proteine anschließend wie bereits für die Wildtyp- $\Delta 29$ NCS geschehen mit Hilfe des ImmunoCAP-Verfahrens (Bousquet *et al.*, 1990) an einer Reihe von Birkenpollenallergiker-Patientenseren auf Kreuzreaktivität mit Bet v 1 getestet werden. Die bereits untersuchten mutierten Proteine $\Delta 29$ NCS_1M und $\Delta 29$ NCS_3M zeigen tatsächlich bei einigen der Patientenseren leicht erhöhte Kreuzreaktivität (Vieths, unveröffentlicht). Die noch ausstehende Untersuchung der bereits konstruierten und eventuell weiterer Mutanten wird zeigen, ob die strukturelle Ähnlichkeit der NCS ausreicht, um in Kombination mit der entsprechenden lokalen Aminosäuresequenz Epitope des Birkenpollenallergens Bet v 1 nachzubilden. Neben der Mutationsanalyse in dem Sinne, dass Epitope des eigentlichen Allergens durch Mutation zerstört werden, der Mimotopanalyse und der Strukturbestimmung von IgG/E-Allergen-Komplexen könnte damit eine weitere aussagekräftige Methode zur Untersuchung von IgE-Epitopen auf der Oberfläche allergener Strukturen etabliert werden.

5 Abkürzungsverzeichnis

3D	dreidimensional
3,4-DHPAA	3,4-Dihydroxyphenylacetaldehyd
4-HPAA	4-Hydroxyphenylacetaldehyd
BIA	Benzylisochinolinalkaloide
CD	Circulardichroismus
CSBP	Cytokinin-bindende Proteine
CSI	<i>chemical shift index</i>
CSP	<i>chemical shift perturbation</i>
<i>E. coli</i>	<i>Escherichia coli</i>
HBA	4-Hydroxyphenylbenzaldehyd
HSQC	<i>heteronuclear single quantum coherence</i>
IgE	Immunoglobuline des Typ E
IgG	Immunoglobuline des Typ G
IL	Interleukin
MHC II	Haupthistokompatibilitätskomplex der Klasse II
MLP	<i>major latex proteins</i>
NCS	(S)-Norcocclaurin-Synthase
NMR	magnetische Kernspinresonanz
OAS	orales Allergie-Syndrom
PR10-Proteine	<i>pathogenesis-related proteins</i> der Klasse 10
RRP	<i>ripening-related proteins</i>
SDS-PAGE	Natriumdodecylsulfat-Polyacrylamid-Gelelektrophorese
<i>Tf</i> -NCS	(S)-Norcocclaurin-Synthase aus <i>Thalictrum flavum</i>
Th-Zellen	T-Helferzellen
TROSY	<i>transverse relaxation optimized spectroscopy</i>

6 Literaturverzeichnis

Altmann F (2007) The role of protein glycosylation in allergy. *Int Arch Allergy Immunol* **142**: 99-115

Bais HP, Vepachedu R, Lawrence CB, Stermitz FR und Vivanco JM (2003) Molecular and biochemical characterization of an enzyme responsible for the formation of hypericin in St. John's wort (*Hypericum perforatum* L.). *J Biol Chem* **278**: 32413-32422

Ballmer-Weber BK, Holzhauser T, Scibilia J, Mittag D, Zisa G, Ortolani C, Oesterballe M, Poulsen LK, Vieths S und Bindslev-Jensen C (2007) Clinical characteristics of soybean allergy in Europe: a double-blind, placebo-controlled food challenge study. *J Allergy Clin Immunol* **119**: 1489-1496

Ballmer-Weber BK, Vieths S (2008) Soy allergy in perspective. *Curr Opin Allergy Clin Immunol* **8**: 270-275

Battersby AR, Binks R (1960) Biosynthesis of Morphine: Formation of Morphine from Norlaudanoline. *Proc Chem Soc*: 360-361

Bousquet J, Chanaz P, Chanal I und Michel FB (1990) Comparison between RAST and Pharmacia CAP system: a new automated specific IgE assay. *J Allergy Clin Immunol* **85**: 1039-1043

Breiteneder H, Mills ENC (2006) Structural bioinformatic approaches to understand cross-reactivity. *Mol Nutr Food Res* **50**: 628-632

Breiteneder H, Mills ENC (2005) Plant food allergens--structural and functional aspects of allergenicity. *Biotechnol Adv* **23**: 395-399

Breiteneder H, Pettenburger K, Bito A, Valenta R, Kraft D, Rumpold H, Scheiner O und Breitenbach M (1989) The gene coding for the major birch pollen allergen Bet v 1, is highly homologous to a pea disease resistance response gene. *EMBO J* **8**: 1935-1938

Buchanan B, Gruissem W, Jones RE (2000) *Biochemistry & Molecular Biology of Plants*. American Society of Plant Physiologists, Rockville, Md.

Bufe A, Spangfort MD, Kahlert H, Schlaak M und Becker WM (1996) The major birch pollen allergen, Bet v 1, shows ribonuclease activity. *Planta* **199**: 413-415

Burks W, Ballmer-Weber BK (2006) Food allergy. *Mol Nutr Food Res* **50**: 595-603

Chapman MD, Pomes A, Breiteneder H und Ferreira F (2007) Nomenclature and structural biology of allergens. *J Allergy Clin Immunol* **119**: 414-420

Chevallier A (2000) *Die BLV-Enzyklopädie der Heilpflanzen*. BLV Verlagsgesellschaft mbH, München

- Chrobok N** (2008) *Epitope grafting* an Norcoclaurin-Synthase: Reinigung und spektroskopische Charakterisierung von Mutanten des Bet v 1-homologen Enzyms. Bachelorarbeit, Lehrstuhl für Struktur und Chemie der Biopolymere, Universität Bayreuth
- Clore GM, Gronenborn AM** (1998) Determining the structures of large proteins and protein complexes by NMR. *Trends Biotechnol* **16**: 22-34
- Crowell DN, John ME, Russell D und Amasino RM** (1992) Characterization of a stress-induced, developmentally regulated gene family from soybean. *Plant Mol Biol* **18**: 459-466
- Davies JM, O'Hehir RE** (2008) Immunogenetic characteristics of immunoglobulin E in allergic disease. *Clin Exp Allergy* **38**: 566-578
- Ebner C, Hoffmann-Sommergruber K und Breiteneder H** (2001) Plant food allergens homologous to pathogenesis-related proteins. *Allergy* **56 Suppl 67**: 43-44
- Esch RE** (2008) Grass pollen allergens. *Clin Allergy Immunol* **21**: 107-126
- Fedorov AA, Ball T, Mahoney NM, Valenta R und Almo SC** (1997) The molecular basis for allergen cross-reactivity: crystal structure and IgE-epitope mapping of birch pollen profilin. *Structure* **5**: 33-45
- Fernandez C, Wider G** (2003) TROSY in NMR studies of the structure and function of large biological macromolecules. *Curr Opin Struct Biol* **13**: 570-580
- Fernandez-Carneado J, Grell D, Durieux P, Hauert J, Kovacsovics T und Tuchscherer G** (2000) Surface grafting onto template-assembled synthetic protein scaffolds in molecular recognition. *Biopolymers* **55**: 451-458
- Ferreira F, Ebner C, Kramer B, Casari G, Briza P, Kungl AJ, Grimm R, Jahn-Schmid B, Breiteneder H, Kraft D, Breitenbach M, Rheinberger HJ und Scheiner O** (1998) Modulation of IgE reactivity of allergens by site-directed mutagenesis: potential use of hypoallergenic variants for immunotherapy. *FASEB J* **12**: 231-242
- Gajhede M, Osmark P, Poulsen FM, Ipsen H, Larsen JN, Joost van Neerven RJ, Schou C, Lowenstein H und Spangfort MD** (1996) X-ray and NMR structure of Bet v 1, the origin of birch pollen allergy. *Nat Struct Biol* **3**: 1040-1045
- Gall H, Kalveram KJ, Forck G und Sterry W** (1994) Kiwi fruit allergy: a new birch pollen-associated food allergy. *J Allergy Clin Immunol* **94**: 70-76
- Gardner KH, Kay LE** (1998) The use of ^2H , ^{13}C , ^{15}N multidimensional NMR to study the structure and dynamics of proteins. *Annu Rev Biophys Biomol Struct* **27**: 357-406
- Garn H, Renz H** (2007) Epidemiological and immunological evidence for the hygiene hypothesis. *Immunobiology* **212**: 441-452
- Garner LA** (2004) Contact dermatitis to metals. *Dermatol Ther* **17**: 321-327
- Gell PGH, Coombs RRA** (1963) Clinical aspects of immunology. Blackwell, Oxford

Gonzalez R, Polo F, Zapatero L, Caravaca F und Carreira J (1992) Purification and characterization of major inhalant allergens from soybean hulls. *Clin Exp Allergy* **22**: 748-755

Gould HJ, Sutton BJ (2008) IgE in allergy and asthma today. *Nat Rev Immunol* **8**: 205-217

Gould HJ, Sutton BJ, Bevil AJ, Bevil RL, McCloskey N, Coker HA, Fear D und Smurthwaite L (2003) The biology of IGE and the basis of allergic disease. *Annu Rev Immunol* **21**: 579-628

Grabley S, Thiericke R (2000) *Drug Discovery from Nature*. Springer, Berlin

Harborne JB, Baxter H, Moss GP (1999) *Phytochemical Dictionary - A Handbook of Bioactive Compounds from Plants*, Ed 2. Taylor & Francis, London

Hayek B, Vangelista L, Pastore A, Sperr WR, Valent P, Vrtala S, Niederberger V, Twardosz A, Kraft D und Valenta R (1998) Molecular and immunologic characterization of a highly cross-reactive two EF-hand calcium-binding alder pollen allergen, Aln g 4: structural basis for calcium-modulated IgE recognition. *J Immunol* **161**: 7031-7039

Heppt W, Bachert C (1998) *Praktische Allergologie*. Thieme, Stuttgart

Hoffmann-Sommergruber K (2000) Plant allergens and pathogenesis-related proteins. What do they have in common? *Int Arch Allergy Immunol* **122**: 155-166

Hoffmann-Sommergruber K, Vanek-Krebitz M, Radauer C, Wen J, Ferreira F, Scheiner O und Breiteneder H (1997) Genomic characterization of members of the Bet v 1 family: genes coding for allergens and pathogenesis-related proteins share intron positions. *Gene* **197**: 91-100

Ilari A, Franceschini S, Bonamore A, Arenghi F, Botta B, Macone A, Pasquo A, Bellucci L und Boffi A (2009) Structural basis of enzymatic (s)-norcoclaurine biosynthesis. *J Biol Chem* **284**: 897-904

Ipsen H, Hansen OC (1991) The NH₂-terminal amino acid sequence of the immunochemically partial identical major allergens of Alder (*Alnus glutinosa*) Aln g I, birch (*Betula verrucosa*) Bet v I, hornbeam (*Carpinus betulus*) Car b I and oak (*Quercus alba*) Que a I pollens. *Mol Immunol* **28**: 1279-1288

Ipsen H, Lowenstein H (1983) Isolation and immunochemical characterization of the major allergen of birch pollen (*Betula verrucosa*). *J Allergy Clin Immunol* **72**: 150-159

Janeway CAj, Travers P, Walport M, Shlomchik M (2002) *Immunologie*, Ed 5. Spektrum, Heidelberg

Karamloo F, Wangorsch A, Kasahara H, Davin LB, Hausteiner D, Lewis NG und Vieths S (2001) Phenylcoumaran benzylic ether and isoflavonoid reductases are a new class of cross-reactive allergens in birch pollen, fruits and vegetables. *Eur J Biochem* **268**: 5310-5320

Kay AB (2001) Allergy and allergic diseases. First of two parts. *N Engl J Med* **344**: 30-37

- Kleine-Tebbe J, Vogel L, Crowell DN, Hausteин UF und Vieths S** (2002) Severe oral allergy syndrome and anaphylactic reactions caused by a Bet v 1- related PR-10 protein in soybean, SAM22. *J Allergy Clin Immunol* **110**: 797-804
- Kutchan TM** (1993) Strictosidine: from alkaloid to enzyme to gene. *Phytochemistry* **32**: 493-506
- Lehmann K, Schweimer K, Reese G, Randow S, Suhr M, Becker WM, Vieths S und Rösch P** (2006) Structure and stability of 2S albumin-type peanut allergens: implications for the severity of peanut allergic reactions. *Biochem J* **395**: 463-472
- Liscombe DK, Facchini PJ** (2008) Evolutionary and cellular webs in benzyloquinoline alkaloid biosynthesis. *Curr Opin Biotechnol* **19**: 173-180
- Liscombe DK, MacLeod BP, Loukanina N, Nandi OI und Facchini PJ** (2005) Evidence for the monophyletic evolution of benzyloquinoline alkaloid biosynthesis in angiosperms. *Phytochemistry* **66**: 2501-2520
- Loeffler S, Stadler R, Nagakura N und Zenk MH** (1987) Norcoclaurine as Biosynthetic Precursor of Thebaine and Morphine. *J Chem Soc* **15**: 1160-1162
- Luk LY, Bunn S, Liscombe DK, Facchini PJ und Tanner ME** (2007) Mechanistic studies on norcoclaurine synthase of benzyloquinoline alkaloid biosynthesis: an enzymatic Pictet-Spengler reaction. *Biochemistry (N Y)* **46**: 10153-10161
- Luttkopf D, Muller U, Skov PS, Ballmer-Weber BK, Wuthrich B, Skamstrup Hansen K, Poulsen LK, Kastner M, Hausteин D und Vieths S** (2002) Comparison of four variants of a major allergen in hazelnut (*Corylus avellana*) Cor a 1.04 with the major hazel pollen allergen Cor a 1.01. *Mol Immunol* **38**: 515-525
- Ma X, Panjikar S, Koepke J, Loris E und Stockigt J** (2006) The structure of Rauvolfia serpentina strictosidine synthase is a novel six-bladed beta-propeller fold in plant proteins. *Plant Cell* **18**: 907-920
- Madigan MT, Martinko JM, Parker J** (2003) Brock Mikrobiologie, Ed 9. Spektrum Akademischer Verlag, Heidelberg
- Markovic-Housley Z, Degano M, Lamba D, von Roepenack-Lahaye E, Clemens S, Susani M, Ferreira F, Scheiner O und Breiteneder H** (2003) Crystal structure of a hypoallergenic isoform of the major birch pollen allergen Bet v 1 and its likely biological function as a plant steroid carrier. *J Mol Biol* **325**: 123-133
- Marsh DG, Goodfriend L, King TP, Lowenstein H und Platts-Mills TA** (1986) Allergen nomenclature. *Bull World Health Organ* **64**: 767-774
- McCoy MA, Wyss DF** (2000) Alignment of weakly interacting molecules to protein surfaces using simulations of chemical shift perturbations. *J Biomol NMR* **18**: 189-198
- Miescher SM, Vogel M** (2002) Molecular aspects of allergy. *Mol Aspects Med* **23**: 413-462
- Mills ENC, Shewry PR** (2004) Plant Food Allergens, Ed 1. Blackwell Science Ltd, Oxford

- Minami H, Dubouzet E, Iwasa K und Sato F** (2007) Functional analysis of norcoclaurine synthase in *Coptis japonica*. *J Biol Chem* **282**: 6274-6282
- Mirza O, Henriksen A, Ipsen H, Larsen JN, Wissenbach M, Spangfort MD und Gajhede M** (2000) Dominant Epitopes and Allergic Cross-Reactivity: Complex formation Between a Fab Fragment of a Monoclonal Murine IgG Antibody and the Major Allergen from Birch Pollen Bet v 1. *J Immunol* **165**: 331-338
- Mittag D, Vieths S, Vogel L, Becker WM, Rihs HP, Helbling A, Wuthrich B und Ballmer-Weber BK** (2004) Soybean allergy in patients allergic to birch pollen: clinical investigation and molecular characterization of allergens. *J Allergy Clin Immunol* **113**: 148-154
- Mogensen JE, Ferreras M, Wimmer R, Petersen SV, Enghild JJ und Otzen DE** (2007) The major allergen from birch tree pollen, Bet v 1, binds and permeabilizes membranes. *Biochemistry* **46**: 3356-3365
- Mogensen JE, Wimmer R, Larsen JN, Spangfort MD und Otzen DE** (2002) The major birch allergen, Bet v 1, shows affinity for a broad spectrum of physiological ligands. *J Biol Chem* **277**: 23684-23692
- Mohapatra SS, Lockey RF und Polo F** (2008) Weed pollen allergens. *Clin Allergy Immunol* **21**: 127-139
- Moiseyev GP, Fedoreyeva LI, Zhuravlev YN, Yasnetskaya E, Jekel PA und Beintema JJ** (1997) Primary structures of two ribonucleases from ginseng calluses. New members of the PR-10 family of intracellular pathogenesis-related plant proteins. *FEBS Lett* **407**: 207-210
- Mothes N, Horak F und Valenta R** (2004) Transition from a botanical to a molecular classification in tree pollen allergy: implications for diagnosis and therapy. *Int Arch Allergy Immunol* **135**: 357-373
- Neudecker P, Lehmann K und Rösch P** (2003) Sequence-specific ¹H, ¹³C and ¹⁵N resonance assignments of SAM22, an allergenic stress-induced protein from soy bean. *J Biomol NMR* **26**: 191-192
- Neudecker P, Nerkamp J, Eisenmann A, Nourse A, Lauber T, Schweimer K, Lehmann K, Schwarzingen S, Ferreira F und Rösch P** (2004) Solution structure, dynamics, and hydrodynamics of the calcium-bound cross-reactive birch pollen allergen Bet v 4 reveal a canonical monomeric two EF-hand assembly with a regulatory function. *J Mol Biol* **336**: 1141-1157
- Neudecker P, Schweimer K, Nerkamp J, Scheurer S, Vieths S, Sticht H und Rösch P** (2001) Allergic cross-reactivity made visible: solution structure of the major cherry allergen Pru av 1. *J Biol Chem* **276**: 22756-22763
- Oberhuber C, Ma Y, Wopfner N, Gadermaier G, Dedic A, Niggemann B, Maderegger B, Gruber P, Ferreira F, Scheiner O und Hoffmann-Sommergruber K** (2008) Prevalence of IgE-binding to Art v 1, Art v 4 and Amb a 1 in mugwort-allergic patients. *Int Arch Allergy Immunol* **145**: 94-101

- Pasquo A, Bonamore A, Franceschini S, Macone A, Boffi A und Ilari A** (2008) Cloning, expression, crystallization and preliminary X-ray data analysis of norcoclaurine synthase from *Thalictrum flavum*. *Acta Crystallogr Sect F Struct Biol Cryst Commun* **64**: 281-283
- Pasternak O, Bujacz GD, Fujimoto Y, Hashimoto Y, Jelen F, Otlewski J, Sikorski MM und Jaskolski M** (2006) Crystal structure of *Vigna radiata* cytokinin-specific binding protein in complex with zeatin. *Plant Cell* **18**: 2622-2634
- Pastorello EA, Farioli L, Pravettoni V, Ispano M, Scibola E, Trambaioli C, Giuffrida MG, Ansaloni R, Godovac-Zimmermann J, Conti A, Fortunato D und Ortolani C** (2000) The maize major allergen, which is responsible for food-induced allergic reactions, is a lipid transfer protein. *J Allergy Clin Immunol* **106**: 744-751
- Pedersen MH, Holzhauser T, Bisson C, Conti A, Jensen LB, Skov PS, Bindslev-Jensen C, Brinch DS und Poulsen LK** (2008) Soybean allergen detection methods--a comparison study. *Mol Nutr Food Res* **52**: 1486-1496
- Pervushin K, Riek R, Wider G und Wuthrich K** (1997) Attenuated T2 relaxation by mutual cancellation of dipole-dipole coupling and chemical shift anisotropy indicates an avenue to NMR structures of very large biological macromolecules in solution. *Proc Natl Acad Sci U S A* **94**: 12366-12371
- Purohit A, Niederberger V, Kronqvist M, Horak F, Gronneberg R, Suck R, Weber B, Fiebig H, van Hage M, Pauli G, Valenta R und Cromwell O** (2008) Clinical effects of immunotherapy with genetically modified recombinant birch pollen Bet v 1 derivatives. *Clin Exp Allergy* **38**: 1514-1525
- Radauer C, Breiteneder H** (2007) Evolutionary biology of plant food allergens. *J Allergy Clin Immunol* **120**: 518-525
- Radauer C, Breiteneder H** (2006) Pollen allergens are restricted to few protein families and show distinct patterns of species distribution. *J Allergy Clin Immunol* **117**: 141-147
- Radauer C, Bublin M, Wagner S, Mari A und Breiteneder H** (2008a) Allergens are distributed into few protein families and possess a restricted number of biochemical functions. *J Allergy Clin Immunol* **121**: 847-52.e7
- Radauer C, Lackner P und Breiteneder H** (2008b) The Bet v 1 fold: an ancient, versatile scaffold for binding of large, hydrophobic ligands. *BMC Evol Biol* **8**: 286
- Rihs HP, Chen Z, Rueff F, Petersen A, Rozynek P, Heimann H und Baur X** (1999) IgE binding of the recombinant allergen soybean profilin (rGly m 3) is mediated by conformational epitopes. *J Allergy Clin Immunol* **104**: 1293-1301
- Ring J, Eberlein-Koenig B und Behrendt H** (2001) Environmental pollution and allergy. *Ann Allergy Asthma Immunol* **87**: 2-6
- Robinson R** (1917) A theory of the mechanism of the phytochemical synthesis of certain alkaloids. *J Chem Soc* **111**: 876-899
- Romagnani S** (2006) Regulation of the T cell response. *Clin Exp Allergy* **36**: 1357-1366

Romagnani S (2004) The increased prevalence of allergy and the hygiene hypothesis: missing immune deviation, reduced immune suppression, or both? *Immunology* **112**: 352-363

Rost B (1999) Twilight zone of protein sequence alignments. *Protein Eng* **12**: 85-94

Rueffer M, El-Shagi H, Nagakura N und Zenk MH (1981) (S)-Norlaudanosoline Synthase: The First Enzyme in the Benzyloquinoline Biosynthetic Pathway. *FEBS Lett* **129**: 5-9

Salzmann M, Pervushin K, Wider G, Senn H und Wuthrich K (1998) TROSY in triple-resonance experiments: new perspectives for sequential NMR assignment of large proteins. *Proc Natl Acad Sci U S A* **95**: 13585-13590

Samanani N, Facchini PJ (2002) Purification and characterization of norcoclaurine synthase. The first committed enzyme in benzyloquinoline alkaloid biosynthesis in plants. *J Biol Chem* **277**: 33878-33883

Samanani N, Facchini PJ (2001) Isolation and partial characterization of norcoclaurine synthase, the first committed step in benzyloquinoline alkaloid biosynthesis, from opium poppy. *Planta* **213**: 898-906

Samanani N, Liscombe DK und Facchini PJ (2004) Molecular cloning and characterization of norcoclaurine synthase, an enzyme catalyzing the first committed step in benzyloquinoline alkaloid biosynthesis. *Plant J* **40**: 302-313

Saraste M, Sibbald PR und Wittinghofer A (1990) The P-loop--a common motif in ATP- and GTP-binding proteins. *Trends Biochem Sci* **15**: 430-434

Schein CH, Ivanciuc O und Braun W (2007) Bioinformatics approaches to classifying allergens and predicting cross-reactivity. *Immunol Allergy Clin North Am* **27**: 1-27

Schirmer T, Hoffmann-Sommergrube K, Susani M, Breiteneder H und Markovic-Housley Z (2005) Crystal structure of the major celery allergen Api g 1: molecular analysis of cross-reactivity. *J Mol Biol* **351**: 1101-1109

Schöning B, Vieths S, Petersen A und Baltes W (1995) Identification and characterisation of allergens related to Bet v 1, the major birch pollen allergen in apple, cherry, celery and carrot by two-dimensional immunoblotting and N-terminal microsequencing. *J Sci Food Agric* **67**: 431-440

Schumacher P (1995) Biophysikalische Therapie der Allergien, Ed 2. Johannes Sonntag Verlagsbuchhandlung GmbH, Stuttgart

Schweimer K, Petersen A, Suck R, Becker WM, Rosch P und Matecko I (2008) Solution structure of Phl p 3, a major allergen from timothy grass pollen. *Biol Chem* **389**: 919-923

Sia SK, Kim PS (2003) Protein grafting of an HIV-1-inhibiting epitope. *Proc Natl Acad Sci U S A* **100**: 9756-9761

Stadler R, Kutchan TM, Löffler S, Nagakura N, Cassels BK und Zenk MH (1987) Revision of the early steps of reticuline biosynthesis. *Tetrahedron Lett* **28**: 1251-1254

- Stadler R, Kutchan TM und Zenk MH** (1989) Norcoclaurine is the central intermediate in benzyloisoquinoline alkaloid biosynthesis. *Phytochemistry* **28**: 1083-1086
- Stadler R, Zenk MH** (1990) A Revision of the Generally Accepted Pathway for the Biosynthesis of the Benzyltetrahydroisoquinoline Alkaloid Reticuline. *Liebigs Ann Chem*: 555-562
- Swoboda I, Jilek A, Ferreira F, Engel E, Hoffmann-Sommergruber K, Scheiner O, Kraft D, Breiteneder H, Pittenauer E und Schmid E** (1995) Isoforms of Bet v 1, the major birch pollen allergen, analyzed by liquid chromatography, mass spectrometry, and cDNA cloning. *J Biol Chem* **270**: 2607-2613
- Swoboda I, Scheiner O, Kraft D, Breitenbach M, Heberle-Bors E und Vicente O** (1994) A birch gene family encoding pollen allergens and pathogenesis-related proteins. *Biochim Biophys Acta* **1219**: 457-464
- Swoboda I, Twaroch T, Valenta R und Grote M** (2008) Tree pollen allergens. *Clin Allergy Immunol* **21**: 87-105
- Takano H, Yanagisawa R und Inoue K** (2007) Components of diesel exhaust particles diversely enhance a variety of respiratory diseases related to infection or allergy: extracted organic chemicals and the residual particles after extraction differently affect respiratory diseases. *J Clin Biochem Nutr* **40**: 101-107
- Tsujishita Y, Hurley JH** (2000) Structure and lipid transport mechanism of a StAR-related domain. *Nat Struct Biol* **7**: 408-414
- Turner H, Kinet JP** (1999) Signalling through the high-affinity IgE receptor Fc epsilonRI. *Nature* **402**: B24-30
- van Loon LC, Rep M und Pieterse CM** (2006) Significance of inducible defense-related proteins in infected plants. *Annu Rev Phytopathol* **44**: 135-162
- van Ree R, Cabanes-Macheteau M, Akkerdaas J, Milazzo JP, Loutelier-Bourhis C, Rayon C, Villalba M, Koppelman S, Aalberse R, Rodriguez R, Faye L und Lerouge P** (2000) Beta(1,2)-xylose and alpha(1,3)-fucose residues have a strong contribution in IgE binding to plant glycoallergens. *J Biol Chem* **275**: 11451-11458
- Venarske D, deShazo RD** (2003) Molecular mechanisms of allergic disease. *South Med J* **96**: 1049-1054
- Verdino P, Barderas R, Villalba M, Westritschnig K, Valenta R, Rodriguez R und Keller W** (2008) Three-dimensional structure of the cross-reactive pollen allergen Che a 3: visualizing cross-reactivity on the molecular surfaces of weed, grass, and tree pollen allergens. *J Immunol* **180**: 2313-2321
- Vieths S, Scheurer S und Ballmer-Weber B** (2002) Current understanding of cross-reactivity of food allergens and pollen. *Ann N Y Acad Sci* **964**: 47-68
- von Pirquet C** (1906) Allergie. *Münchener Medizinische Wochenschrift* **53**: 1547-1548

Westphal S, Kolarich D, Foetisch K, Lauer I, Altmann F, Conti A, Crespo JF, Rodriguez J, Enrique E, Vieths S und Scheurer S (2003) Molecular characterization and allergenic activity of Lyc e 2 (beta-fructofuranosidase), a glycosylated allergen of tomato. *Eur J Biochem* **270**: 1327-1337

Wills-Karp M, Santeliz J und Karp CL (2001) The germless theory of allergic disease: revisiting the hygiene hypothesis. *Nat Rev Immunol* **1**: 69-75

Wilson ML, Coscia CJ (1975) Studies on the Early Stages of *Papaver* Alkaloid Biogenesis. *J Am Chem Soc* **97**: 431-432

Wink M (1999a) Biochemistry of plant secondary metabolism. *Annual plant reviews*, Volume 2, Sheffield Academic Press, Sheffield

Wink M (1999b) Functions of Plant Secondary Metabolites and their Exploitation in Biotechnology. *Annual plant reviews*, Volume 3, Sheffield Academic Press, Sheffield

Winterstein E, Trier G (1910) *Die Alkaloide*. Borntraeger, Berlin

Ziegler J, Facchini PJ (2008) Alkaloid biosynthesis: metabolism and trafficking. *Annu Rev Plant Biol* **59**: 735-769

Zuiderweg ER (2002) Mapping protein-protein interactions in solution by NMR spectroscopy. *Biochemistry* **41**: 1-7

7 Publikationsliste

7.1 Einzelarbeit A

Hanna Berkner*, Philipp Neudecker*, Diana Mittag, Barbara K. Ballmer-Weber, Kristian Schweimer, Stefan Vieths und Paul Rösch (2008): Cross-reactivity of pollen and food allergens: Soybean Gly m 4 is a member of the Bet v 1 superfamily and closely resembles yellow lupine proteins. *Bioscience Reports*, doi:10.1042/BSR20080117

*beide Autoren haben im gleichen Maße zur Arbeit beigetragen

Mit Ausnahme des NMR-Experimentes zur Bestimmung des heteronuklearen NOE, das von Kristian Schweimer und mir aufgenommen wurde, wurden alle NMR-Experimente von Philipp Neudecker durchgeführt. Die Spektren wurden von Philipp Neudecker und mir mit Hilfe von Kristian Schweimer ausgewertet, und die Strukturrechnungen wurden von Philipp Neudecker, Kristian Schweimer und mir durchgeführt. Die immunologischen Experimente wurden von Diana Mittag unter der Leitung von Barbara Ballmer-Weber und Stefan Vieths durchgeführt und ausgewertet. Das Manuskript wurde von Paul Rösch, Kristian Schweimer und mir zusammen mit Diana Mittag und Stefan Vieths verfasst.

7.2 Einzelarbeit B

Hanna Berkner, Julia Engelhorn, David K. Liscombe, Kristian Schweimer, Birgitta M. Wöhlrl, Peter J. Facchini, Paul Rösch und Irena Matečko (2007): High-yield expression and purification of isotopically labeled norcoclaurine synthase, a Bet v 1-homologous enzyme, from *Thalictrum flavum* for NMR studies. *Protein Expression and Purification* **56**(2), 197-204

Die Reinigung des ursprünglichen von Peter J. Facchini zur Verfügung gestellten *Tf*-NCS-Konstruktes wurde von Julia Engelhorn unter Anleitung von Irena Matečko durchgeführt. Die Optimierung der Expression und Reinigung wurde von mir mit Unterstützung von Irena Matečko und Birgitta M. Wöhlrl geplant und experimentell von mir durchgeführt. Die Enzymaktivität wurde von David K. Liscombe unter Leitung von Peter J. Facchini getestet.

Ich habe die NMR-Spektren aufgenommen und mit Hilfe von Kristian Schweimer ausgewertet. Das Manuskript wurde von mir zusammen mit Irena Matečko und Paul Rösch verfasst.

7.3 Einzelarbeit C

Hanna Berkner, Kristian Schweimer, Irena Matečko und Paul Rösch (2008): Conformation, catalytic site, and enzymatic mechanism of the PR10 allergen-related enzyme norcoclaurine synthase. *Biochemical Journal* **413**(2), 281-290

Ich habe die CD-spektroskopischen Messungen und die analytische Größenausschluss-Chromatographie durchgeführt und ausgewertet. Alle NMR-Experimente wurden von mir mit Unterstützung von Irena Matečko und Kristian Schweimer gemessen und von mir analysiert. Das Homologiemodell der *Tf*-NCS wurde von Irena Matečko, Kristian Schweimer und mir konstruiert. Das Manuskript wurde von mir zusammen mit Irena Matečko, Kristian Schweimer und Paul Rösch verfasst.

8 Einzelarbeiten

Einzelarbeit A

Hanna Berkner*, Philipp Neudecker*, Diana Mittag, Barbara K. Ballmer-Weber, Kristian Schweimer, Stefan Vieths und Paul Rösch (2008): Cross-reactivity of pollen and food allergens: Soybean Gly m 4 is a member of the Bet v 1 superfamily and closely resembles yellow lupine proteins. *Bioscience Reports*, doi:10.1042/BSR20080117

*beide Autoren haben im gleichen Maße zur Arbeit beigetragen



Cross-reactivity of pollen and food allergens: soya-bean Gly m 4 is a member of the Bet v 1 superfamily and closely resembles yellow lupine proteins

Hanna BERKNER^{*1}, Philipp NEUDECKER^{*1,2}, Diana MITTAG^{†3}, Barbara K. BALLMER-WEBER[†], Kristian SCHWEIMER^{*}, Stefan VIETHS[‡] and Paul RÖSCH^{*4}

^{*}Department of Biopolymers and Research Center for Bio-Macromolecules, Universität Bayreuth, Universitätsstrasse 30, 95447 Bayreuth, Germany, [†]Allergy Unit, Department of Dermatology, University Hospital Zurich, Gloriastr. 31, 8091 Zurich, Switzerland, and [‡]Department of Allergology, Paul-Ehrlich-Institut, Paul-Ehrlich-Strasse 51-59, 63225 Langen, Germany

Synopsis

In many cases, patients allergic to birch pollen also show allergic reactions after ingestion of certain fruits or vegetables. This observation is explained on the molecular level by cross-reactivity of IgE antibodies induced by sensitization to the major birch pollen allergen Bet v 1 with homologous food allergens. As IgE antibodies recognize conformational epitopes, a precise structural characterization of the allergens involved is necessary to understand cross-reactivity and thus to develop new methods of allergen-specific immunotherapy for allergic patients. Here, we report the three-dimensional solution structure of the soya-bean allergen Gly m 4, a member of the superfamily of Bet v 1 homologous proteins and a cross-reactant with IgE antibodies originally raised against Bet v 1 as shown by immunoblot inhibition and histamine release assays. Although the overall fold of Gly m 4 is very similar to that of Bet v 1, the three-dimensional structure of these proteins differ in detail. The Gly m 4 local structures that display those differences are also found in proteins from yellow lupine with known physiological function. The three-dimensional structure of Gly m 4 may thus shed some light on the physiological function of this subgroup of PR10 proteins (class 10 of pathogenesis-related proteins) and, in combination with immunological data, allow us to propose surface patches that might represent cross-reactive epitopes.

Key words: allergen, epitope, histamine release, NMR, protein structure, soya bean

INTRODUCTION

15–20% of the population of industrialized countries, with tendency to increasing numbers, suffer from pollinosis, which is often caused by sensitization to birch pollen in moderate climate zones of Europe [1]. In approx. 90% of birch pollen-allergic patients, the major birch pollen allergen Bet v 1 is responsible for allergic symptoms such as hay fever or asthma [2]. Additionally, birch pollen allergy is often accompanied by allergic reactions after ingestion of specific plant-derived food, including fruits, nuts and vegetables, and this association finds its molecular correlate

in structural similarities between respective allergens. A type 1 allergic reaction like hay fever is mediated by cross-linking of allergen-specific IgE antibodies bound to mast cells or basophils by allergens, which, in turn, leads to degranulation of effector cells and release of histamine and other mediators provoking allergic symptoms. Allergens that show sufficiently similar IgE epitopes may react with the same IgE molecules. IgE epitopes are thought to be conformational rather than sequential [3], and patients sensitized to one allergen may show allergic reactions to proteins with similar three-dimensional structures. Interestingly, patients in birch pollen-related food allergy frequently show multiple food allergies to several but not all foods known to be

Abbreviations used: HMQC, heteronuclear single quantum correlation; HSQC, heteronuclear single-quantum coherence; NOE, nuclear Overhauser effect; OAS, oral allergy syndrome; PR protein, pathogenesis-related protein; PR10 protein, class 10 of PR protein; RDC, residual dipolar coupling; rmsd, root mean square deviation; SAM22, starvation-associated message 22; TBS, Tris-buffered saline

¹These authors contributed equally to this work.

²Present address: Departments of Molecular Genetics, Biochemistry and Chemistry, University of Toronto, 1 King's College Circle, Toronto, ON, Canada M5S 1A8.

³Present address: The Walter and Eliza Hall Institute of Medical Research, 1G Royal Parade, Parkville, VIC 3050, Australia.

⁴To whom correspondence should be addressed (email roesch@unibt.de).



involved in this clinical syndrome, indicating that differences in the epitope structure of the causative allergen molecules may exist.

The identification of proteins highly homologous with the major birch pollen allergen Bet v 1 in pollen allergy-related foods like cherry, apple, hazelnut, carrot, celery or soya bean (Figure 1) was the initial step towards an explanation for allergic cross-reactivity on the molecular level.

According to their sequence similarities and related joined characteristics (molecular mass: approx. 17 kDa; acidic pI: approx. 6), these proteins are members of the Bet v 1 protein superfamily and the family of PR10 proteins [class 10 of PR proteins (pathogenesis-related proteins)] according to the structural classification of proteins [4]. The question concerning the physiological function of PR10 proteins is still unanswered. Recently, this open question was addressed with the discovery that lupine protein LIPR10.2B serves as a cytokinin reservoir [5] and with a structural study suggesting the norcoclaurin synthase to be a true member of the PR10 family [6]. Clearly, detailed structural studies of the PR10 proteins may contribute to unravel their relations and thus add to the knowledge of functional features of this class of proteins.

The notion that similarity of the three-dimensional structures of allergens is key to understanding cross-reactivity was confirmed by comparison of the conformations of allergens from various plants: the three-dimensional structures of the major birch pollen allergen Bet v 1 [7], the major cherry allergen Pru av 1 [8] and the celery allergen Api g 1 [9] not only comprise identical arrangements of secondary structure elements, but also they are virtually identical, explaining the phenomenon of cross-reactivity on a molecular level. Detailed analysis of common features and differences regarding primary and tertiary structures of Bet v 1 homologous allergens can thus help us to identify specific and cross-reactive epitopes on the molecular surface. Although mutational epitope analysis and mimotope selection can furnish information on putative IgE epitopes of Bet v 1 and its homologues, the complete picture necessary to construct hypoallergenic protein variants suitable for immunotherapy requires structural information in atomic detail.

The Bet v 1 homologous protein Gly m 4 from soya bean was identified on the mRNA level in 1992 [10], and the putative protein was shown to be induced under stress conditions like starvation, appropriating the protein nomenclature SAM22 (starvation-associated message 22). Later, SAM22 was suggested to be the molecule responsible for allergic reactions to soya bean in birch pollen-allergic patients sensitized to Bet v 1 [11], which could be confirmed in subsequent clinical and immunological studies [12]. In consequence, the protein was renamed Gly m 4 according to the general rules for allergen nomenclature [13]. Remarkably, in contrast with other Bet v 1 homologous proteins, the allergic reactions caused by Gly m 4 tend to be severe [14].

As food containing soy protein becomes more and more popular worldwide, soy allergy in patients sensitized to birch pollen represents a major health problem, rendering the detailed structural analysis of Gly m 4 and its immunological properties an important step towards understanding the molecular basis of this

phenomenon and towards the development of safe and efficacious immunotherapy of pollen-related food allergy.

EXPERIMENTAL

Structure determination of Gly m 4

Sample preparation

Recombinant Gly m 4 was expressed, purified and stored as freeze-dried powder as described in [15]. NMR samples were prepared by dissolving either ^{15}N - or ^{15}N , ^{13}C -labelled Gly m 4 in 10–50 mM potassium phosphate buffer (pH 7.0), 10% $^2\text{H}_2\text{O}$ and 0.02% sodium azide. For specific measurement of NOEs (nuclear Overhauser effects) involving aromatic protons, unlabelled Gly m 4 was dissolved in 50 mM potassium phosphate buffer (pH 7.0) in $^2\text{H}_2\text{O}$. The final sample concentrations were 0.7–1.8 mM.

NMR experiments

All NMR experiments were performed at 298 K on Bruker DRX 600 MHz, AV 600 MHz and DMX 750 MHz spectrometers with pulsed field-gradient capabilities according to standard experimental procedures [16]. In addition to the experiments for resonance assignment [15], the following NOESY experiments were performed: two-dimensional [^1H , ^1H]-NOESY using unlabelled Gly m 4 in $^2\text{H}_2\text{O}$, ^{15}N -NOESY-HSQC (heteronuclear single-quantum coherence), ^{13}C -NOESY HSQC with 120 ms mixing time, and ^{15}N , ^{15}N -HMQC (heteronuclear single quantum correlation)-NOESY-HSQC, ^{13}C , ^{15}N -HMQC-NOESY-HSQC with 150 ms mixing time. Slowly exchanging amide protons involved in H-bond formation were identified from a ^{15}N -HSQC spectrum recorded 1 h after the freeze-dried protein was dissolved in $^2\text{H}_2\text{O}$. $^3\text{J}/\text{HNH}\alpha$ scalar coupling constants were derived from the HNHA experiment. $^1\text{D}_{\text{HN}}$ RDC (residual dipolar coupling) constants were measured with a Gly m 4 sample weakly aligned by addition of 18 mg/ml Pf1 phage particles using IPAP (inphase antiphase) experiments. $\{^1\text{H}\}^{15}\text{N}$ steady state NOE values were measured applying standard pulse sequences [17]. NMR data were processed with in-house written software and analysed using NMRView [18].

Structure calculation

NOESY cross peaks were classified according to their relative intensities and converted into distance restraints with upper limits of 3.0 Å (1 Å = 0.1 nm) (strong), 4.0 Å (medium), 5.0 Å (weak) and 6.0 Å (very weak). For ambiguous distance restraints, the r^{-6} summation over all assigned possibilities defined the upper limit. Distances introduced to account for the involvement of an amide proton in an H-bond were set to 2.3 Å between acceptor and amide proton, and 3.3 Å between acceptor and amide nitrogen. Dihedral restraints for Φ angles were derived from $^3\text{J}/\text{HNH}\alpha$ scalar coupling constants including a Φ angle between -80° and -40° for coupling constants smaller than 6.0 Hz, whereas values greater than 8.0 Hz were converted into a

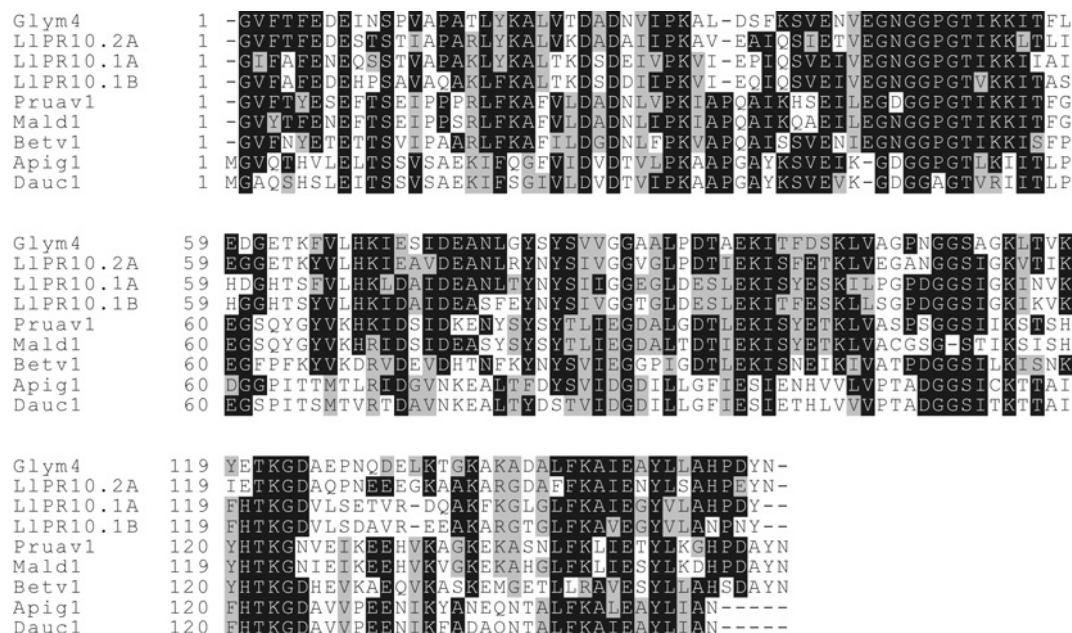


Figure 1 Sequence alignment of Gly m 4 with three homologous proteins from yellow lupine with known structure and the Bet v 1 allergens used in immunoblot inhibition experiments

Sequences are sorted according to their similarity to Gly m 4; black background: identical amino acid; grey background: conservative exchange; threshold for shading: consensus in more than four sequences; alignment performed with ClustalW [43] and shaded with Boxshade 3.21 (<http://www.ch.embnet.org>).

Φ angle between -160° and -80° . Additional dihedral restraints for Φ and ψ angles were derived from chemical shifts using the TALOS program [19] (Table 1). Structure calculations were performed with the program XPLOR-NIH 1.2.1 [20] using a three-step simulated annealing protocol with floating assignment of prochiral groups including a conformational database potential. The 25 structures with the lowest total energy were then refined using the RDCs with a harmonic potential. RDCs measured for amino acids located in areas flexible on the pico- to nanosecond timescale ($\{^1\text{H}\}^{15}\text{N}$ -NOE < 0.65) were excluded from the calculations. The 20 structures showing the lowest values of the target function excluding the database potential were further analysed with the programs XPLOR [20], PROCHECK 3.5.4 [21] and MOLMOL [22] and visualized with PyMOL (DeLano Scientific; <http://pymol.sourceforge.net/>). Surface areas of putative cross-reactive IgE epitopes were calculated using the software nACCESS with a probe radius of 2.5 Å [23]. Co-ordinates were deposited in the PDB (accession code 2K7H).

Soya-bean allergic patients

Sera of six patients (age: 22 ± 5 ; 5 females and 1 male) with food allergy to soy as confirmed by double-blind, placebo-controlled food challenge in the Allergy Unit at the University Hospital Zürich, Switzerland [12,24], were used in the present study. All patients reacted to ingestion of soy with an OAS (oral allergy syndrome) involving symptoms like itching and swelling of the oral cavity. In addition, more systemic symptoms such as urticaria and throat tightness occurred in two patients. All patients had

concomitant OAS to apple, cherry, carrot, celery (except patient I) and peanut.

Immunological characterization of Gly m 4

Histamine release

The histamine release assay was performed using the Histamine Enzyme Immunoassay kit (Immunotech, Marseilles, France) according to the manufacturer's instructions. Whole peripheral blood cells of soya-bean allergic subjects were stimulated after dilution in histamine-release buffer (1:7) provided in the kit. Diluted blood (200 μl) was incubated with 100 μl of different dilutions of allergens (0.0001–100 $\mu\text{g}/\text{ml}$) in PBS for 30 min at 37°C . After centrifugation at 900 g for 5 min at 5°C , the released histamine was measured in the supernatant using the Histamine Enzyme Immunoassay kit. Total histamine was determined after cell lysis by freeze-thawing three times. For analysis of spontaneous histamine release, cells were incubated with PBS instead of diluted allergen.

Immunoblot inhibition

SDS/PAGE [25] was performed using a 15% separation and a 5% stacking gel. rGly m 4 was reduced by heating with dithiothreitol (Sigma–Aldrich, Deisenhofen, Germany), loaded on to the gel at a concentration of 1 $\mu\text{g}/\text{cm}$, and transferred on to 0.2 μm pore-sized nitrocellulose membranes by semi-dry blotting using a Novablot apparatus (Amersham Biosciences, Uppsala, Sweden). Remaining binding sites were blocked twice by incubation with

**Table 1 Structure calculation statistics**

Parameter	Value
Experimental restraints	
Distance restraints	
NOE distance restraints	
Total	2481
Intraresidual	786
Sequential	674
Medium range ($2 \leq \Delta i \leq 4$)	331
Long range ($\Delta i \geq 5$)	673
Ambiguous	17
Hydrogen bonds*	24
Dihedral restraints	
Scalar couplings [$^3J(\text{HN}, \text{H}\alpha)$]	64
TALOS	228
Residual dipolar couplings ($1D_{\text{HN}}$)	109
Properties of the accepted structure ensemble	
Restraint violation	
Average distance restraint violation	$0.004 \text{ \AA} \pm 0.001 \text{ \AA}$
Maximum distance restraint violation	0.24 \AA
Average dihedral restraint violation	$0.47^\circ \pm 0.05^\circ$
Maximum dihedral restraint violation	7.3°
Average RDC restraint violation	$0.14 \text{ Hz} \pm 0.02 \text{ Hz}$
Maximum RDC restraint violation	1.35 Hz
Deviation from ideal geometry	
Bond length	$0.00050 \text{ \AA} \pm 0.00005 \text{ \AA}$
Bond angle	$0.108^\circ \pm 0.007^\circ$
Co-ordinate precision	
Overall rmsd	
Backbone	0.60 \AA
Heavy atoms	0.88 \AA
Secondary structure rmsd amino acids 2–11; 15–23; 26–33; 37–44; 52–59; 62–74; 79–87; 97–105; 111–120; 129–152	
Backbone	0.53 \AA
Heavy atoms	0.83 \AA
Ramachandran plot statistics	
Most favoured regions	92.9%
Additional allowed regions	7.1%
Generously allowed regions	0.0%
Disallowed regions	0.0%

*Two distance restraints per hydrogen bond ($N_{\text{donor}} - O_{\text{acceptor}} \leq 3.3 \text{ \AA}$, $\text{HN}_{\text{donor}} - O_{\text{acceptor}} \leq 2.3 \text{ \AA}$).

0.3% Tween 20 in -TBS (Tris-buffered saline; 50 mM Tris/HCl and 150 mM NaCl, pH 7.4). The nitrocellulose strips were incubated overnight with 130 μl of serum of four patients diluted to 1 ml in incubation buffer (TBS containing 0.05% Tween 20 and 0.1% BSA). For inhibition, strips were incubated with 130 μl of serum and additionally 20 μg of rGly m 4, rBet v 1, rApi g 1, rDau c 1, rPru av 1 and rMal d 1 respectively.

RESULTS AND DISCUSSION

Allergenic activity of rGly m 4 evaluated by histamine release and immunoblot inhibition

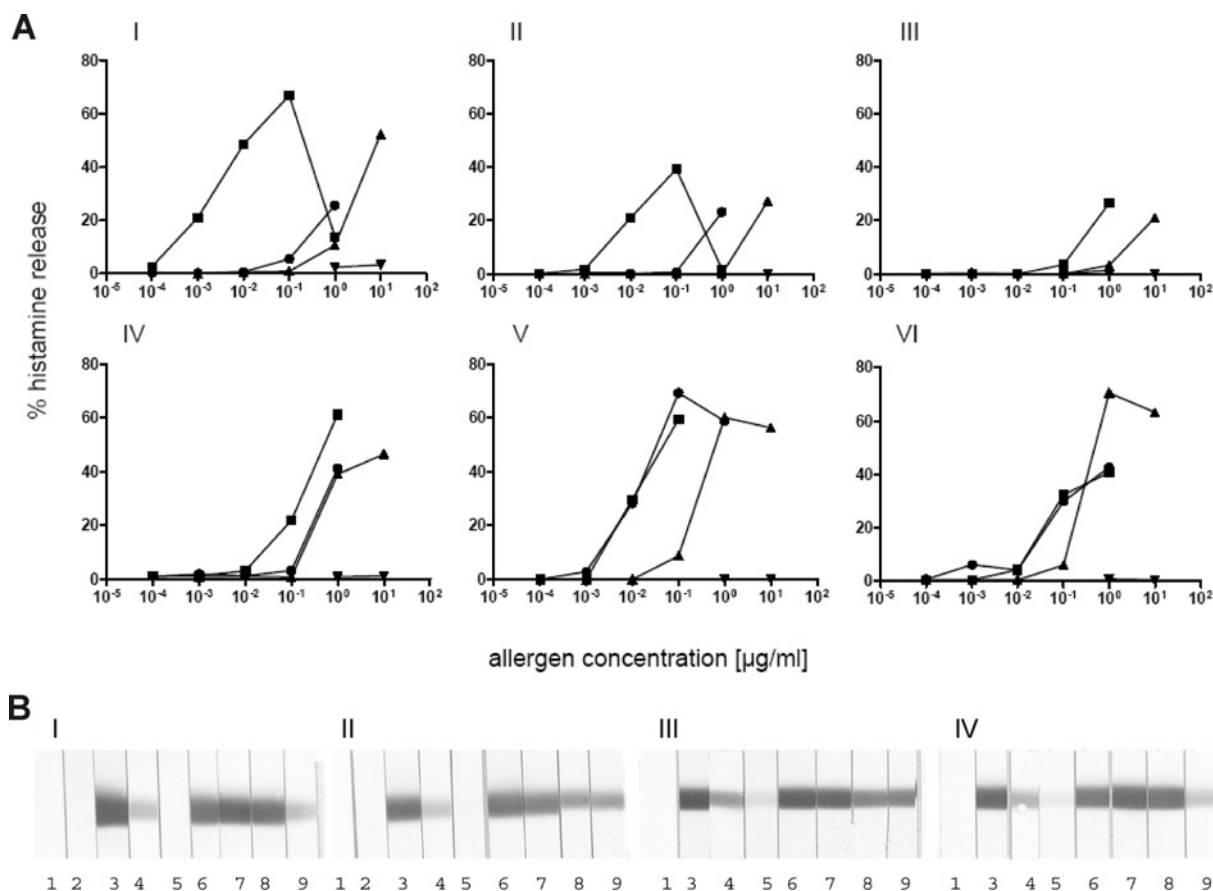
Histamine release (maximum 23–69%) from blood cells after stimulation with rGly m 4 confirmed the biological activity of the recombinant allergen in 5/6 soy allergic subjects (Figure 2A). The similarity between the curves resulting from stimulation with soy isolate and those resulting from stimulation with rGly m 4 clearly indicated that Gly m 4 was the protein responsible for soy allergy in these patients and it proved the equivalence of rGly m 4 as used in the present study and authentic Gly m 4 occurring in soya-bean extract. Bet v 1 elicited histamine release at 10–100-fold lower protein concentrations than Gly m 4, indicating Bet v 1 to be a more potent allergen in these patients. These findings are consistent with Bet v 1 being the primary sensitizing allergen in patients with birch pollen-associated soy allergy.

In immunoblot inhibition, Bet v 1 was found to be the most potent inhibitor of IgE binding to Gly m 4, followed by Gly m 4 and Mal d 1 (Figure 2B). Bet v 1 showed complete inhibition of IgE to Gly m 4 in 4/4 patients that were tested (Figure 2B), whereas Gly m 4 itself showed strong partial inhibition. This finding is consistent with the view that Bet v 1 is the sensitizing allergen. Mal d 1 showed similar strong partial inhibition in all patients. Pru av 1 inhibited IgE binding to Gly m 4 in 2/4 patients (II and III), whereas no inhibition was observed for Api g 1 and Dau c 1 in either patient tested. This difference may be explained by the substitution of the crucial Glu⁴⁴/Glu⁴⁵ in Gly m 4, Bet v 1, Mal d 1 and Pru av 1 for Lys⁴⁴ in Api g 1 and Dau c 1 [26,27].

Solution structure of Gly m 4

Complete backbone amide resonance assignment and assignment of most side chain resonances [15] was the basis for the derivation of 2906 experimental restraints used for structure calculation (Table 1). The resulting ensemble of 20 structures (Figure 3A) is well defined with an overall backbone rmsd (root mean square deviation) of 0.60 \AA and only insignificant violations of experimental restraints. Analysis of the stereochemical properties with PROCHECK 3.5.4 [21] confirms good stereochemical quality of the structure ensemble as mirrored by the fact that more than 90% of the amino acid residues are located in the most favoured regions of the Ramachandran plot.

The solution structure of Gly m 4 is composed of a seven-stranded antiparallel β -sheet and two short V-shaped α -helices $\alpha 1$ and $\alpha 2$ that pack together against the long C-terminal helix $\alpha 3$. The β -sheet consists of strands $2 \rightarrow 11$ ($\beta 1$), $120 \rightarrow 111$ ($\beta 7$), $97 \rightarrow 105$ ($\beta 6$), $87 \rightarrow 79$ ($\beta 5$), $62 \rightarrow 74$ ($\beta 4$), $59 \rightarrow 52$ ($\beta 3$), and $37 \rightarrow 44$ ($\beta 2$). Three of the strands contain β -bulges, namely strand $\beta 2$ at residues Phe³⁷/Lys³⁸ and Val⁴³/Glu⁴⁴, strand $\beta 4$ at Glu⁷¹, and strand $\beta 5$ at Val⁸⁵. Strand $\beta 6$ is interrupted at Leu¹⁰³. The consequence of these irregularities of the β -strands is curving of the whole β -sheet around helix $\alpha 3$ that spans 24 residues from Gln¹²⁹ to Ala¹⁵². Helices $\alpha 1$ and $\alpha 2$ comprise residues Pro¹⁵–Val²³ and Ala²⁶–Ala³³ respectively (Figure 3B).



Q5 **Figure 2** ????

(A) Histamine release in six different patients (i-vi) allergic to soy after stimulation with soy isolate (▲), Gly m 4 (●), Bet v 1 (■) and ovalbumin (▼) as a negative control. (B) Inhibition of IgE binding to Gly m 4 with allergens of the Bet v 1 family in the four patients I-IV with patients II and IV corresponding to patients II and IV in (A): lane 1, no serum; lane 2, serum of non-allergic control patient; lane 3, no inhibitor; lane 4, rGly m 4; lane 5, rBet v 1; lane 6, rApi g 1; lane 7, rDau c 1; lane 8, rPru av 1; lane 9, rMal d 1.

Comparison with Bet v 1 and other homologues

The conformations of three allergenic proteins of the PR10 family are known: Bet v 1 [7], the major cherry allergen Pru av 1 [8] and the celery allergen Api g 1 [9]. Also known are three structures of Bet v 1 homologous proteins from yellow lupine, LIPR10.1A, LIPR10.1B and LIPR10.2A [28,29]. Gly m 4, Bet v 1 and its homologues feature identical secondary-structure elements in identical order. As already evident from the homology model [15], the overall Gly m 4 fold is virtually identical with that of Bet v 1 and includes the hydrophobic cavity formed by the β -sheet and helix $\alpha 3$ that has been suggested to be important for the physiological function of the Bet v 1 family of proteins [8,30]. Superposition of the backbone of the actual experimentally determined Gly m 4 secondary structure elements, however, yields a surprisingly high rmsd of 2.60 Å for Bet v 1, 1.95 Å for Pru av 1 and 3.98 Å for Api g 1 (Figure 5A). The high rmsd values originate from differences in the connection of strands $\beta 3$ and $\beta 4$, in helix $\alpha 3$ and the position of helix $\alpha 3$ relative to the β -sheet.

Whereas amino acids Phe⁵⁸ to Lys⁶⁵ form a flexible loop in Bet v 1, a short type II β -turn, Glu⁵⁹ to Glu⁶², connects the extended strands $\beta 3$ and $\beta 4$ in Gly m 4 as evidenced by the H-bond between the backbone amide proton of Glu⁶² and the backbone carbonyl group of Glu⁵⁹. Such a β -hairpin conformation at the identical position was also identified in three highly Bet v 1 homologous proteins from yellow lupine with the same overall fold [28,29].

The long C-terminal helix $\alpha 3$ of Gly m 4 is bent as in the lupine proteins, contrasting the straight helix $\alpha 3$ in Bet v 1. Whereas in the yellow lupine protein LIPR10.1A, the N-terminal part of the helix is disordered and in LIPR10.2A, a kink of almost 60° is found near Phe¹⁴², the whole helix is well ordered in LIPR10.1B as it is in Gly m 4 [28,29]. The similarity of the helix conformation in LIPR10.1B and Gly m 4 is illustrated by the relatively low backbone rmsd of 0.77 Å, resulting from a comparison of helices $\alpha 3$ exclusively. The same comparison for Gly m 4 and Bet v 1, in contrast, yields an rmsd of 1.16 Å.

Helix $\alpha 3$ of Gly m 4 is shifted relative to the β -sheet by almost one turn along the helix axis as compared with Bet v 1.

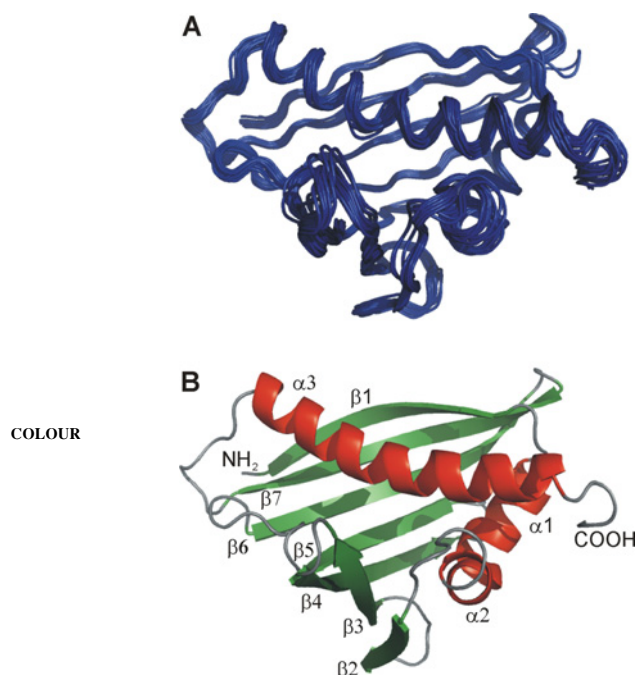


Figure 3 ?????

(A) Backbone overlay of the 20 lowest energy solution structures of Gly m 4; backbone rmsd = 0.6 Å. (B) Cartoon of the average structure of Gly m 4. The bent and shifted C-terminal helix sets this structure slightly apart from the strongly related Bet v 1 structure.

The most important contact points of the helix and the β -sheet are aromatic amino acids with side chains large enough to bridge these structural elements. For example, Phe¹⁴³ contacts Ser¹¹, and Phe³/Phe⁵ both pack against Leu¹³². A corresponding shift of the helix relative to the β -sheet is also observed in the three yellow lupine proteins LIPR10.1A, LIPR10.1B and LIPR10.2A. This results in a comparatively low backbone rmsd value of 1.72 Å for superposition of all secondary structure elements of LIPR10.1B and Gly m 4 as compared with superposition with the Bet v 1 allergens (Figures 5A and 5B).

The loop connecting $\alpha 2$ and $\beta 2$ is shorter in Gly m 4 than it is in Bet v 1, thus adopting a different conformation. Flexibility in terms of high-frequency motion on the picosecond to nanosecond timescale as indicated by reduced $\{^1\text{H}\}^{15}\text{N}$ -NOE values is observed for the N-terminal part of $\beta 7$ near the glycine-rich loop $\beta 6$ – $\beta 7$ and for the extended loop connecting strand $\beta 7$ with helix $\alpha 3$ (Figure 4). The latter was also found to be flexible in Pru av 1 [8]. The glycine-rich loop between $\beta 2$ and $\beta 3$ with the consensus sequence Gly-Xaa-Gly-Gly-Xaa-Gly-Thr ('P-loop') is highly conserved in the Bet v 1 homologous allergens and the lupine proteins (Figure 1). It adopts the same conformation in all these proteins as reflected by a very low rmsd of 0.52 Å for the alignment of all heavy atoms of the P-loop in Bet v 1 and Gly m 4 (Figure 5C). This loop and the neighbouring amino acids, especially the preceding Glu⁴⁴/Glu⁴⁵, have been reported to be important for IgE binding to Bet v 1 and Pru av 1 [8,31,32].

Whereas this glutamate is not conserved in Api g 1 and Dau c 1, it is also present in Gly m 4, and its side chain points in the same direction as in Bet v 1 and LIPR10.1B (Figure 5C).

Links between immunological cross-reactivity and structural properties

Characterizing individual and cross-reactive IgE epitopes of allergens is an important prerequisite for understanding the molecular basis of the clinical phenomenon of pollen-related food allergy and for the design of hypo-allergenic proteins for a safer allergen-specific immunotherapy. A number of studies using techniques like mutational epitope analysis and mimotope selection have been carried out to localize putative epitopes on the surface of Bet v 1 and its homologous tree pollen and food allergens. Surface patches conserved with respect to conformation and charge distribution on homologous allergens are prime candidates for cross-reactive epitopes.

The histamine release experiments revealed a drastically higher allergenic activity of Bet v 1 in comparison with Gly m 4. Consistent with this finding, immunoblot inhibition showed that Bet v 1 had the strongest capacity to inhibit IgE binding to Gly m 4 (i.e. cross-reactivity), probably because Bet v 1 is the sensitizing allergen in pollen-related soy allergy, that is IgE antibodies are primarily produced against the major birch pollen allergen. Mal d 1 also showed significant inhibitory capacity in the immunoblotting experiments, followed by Pru av 1 (Figure 2). No cross-reactivity was observed for Api g 1 and Dau c 1. Thus we searched for surface patches on Gly m 4 consisting of amino acids conserved either only in Bet v 1 or additionally in Mal d 1 and Pru av 1 (Figures 6A and 6B). With this strategy, we could identify three compact conserved surface patches as well as one less coherent patch on Gly m 4 as candidate epitopes (Figure 6C).

One patch consists of the continuous stretch of amino acids Ser³⁹–Ile⁵⁵, corresponding to the area around the P-loop characteristic for the Bet v 1 protein family. This area spans a surface of $\sim 1000 \text{ Å}^2$ in Gly m 4 and has already been identified as a potential cross-reactive epitope in Bet v 1 and Pru av 1 by site-directed mutagenesis [31–33]. The crystal structure of an IgG antibody fragment inhibiting IgE binding in complex with Bet v 1 supports these findings [34]. Especially the glutamate preceding the loop has turned out to be important for IgE binding, as mutation of this residue strongly influences the IgE binding capacity of Bet v 1 as well as that of Pru av 1. The conformation of this loop in Gly m 4 is virtually identical with the Bet v 1/Mal d 1 loops but contrasts the Api g 1/Dau c 1 loops, the latter containing a positively charged lysine residue in the position corresponding to the glutamate. Thus the P-loop region is not only potentially responsible for cross-reactivity of Bet v 1 and Pru av 1, but it may also represent a cross-reactive epitope of Gly m 4 and probably of Mal d 1.

The second potential cross-reactive epitope consists of residues located in the two short helices $\alpha 1$ and $\alpha 2$ and the C-terminus of the protein including parts of helix $\alpha 3$. These residues also form a coherent surface area of $\sim 1100 \text{ Å}^2$, large enough to serve as an IgE-binding site. The structural similarity of Gly

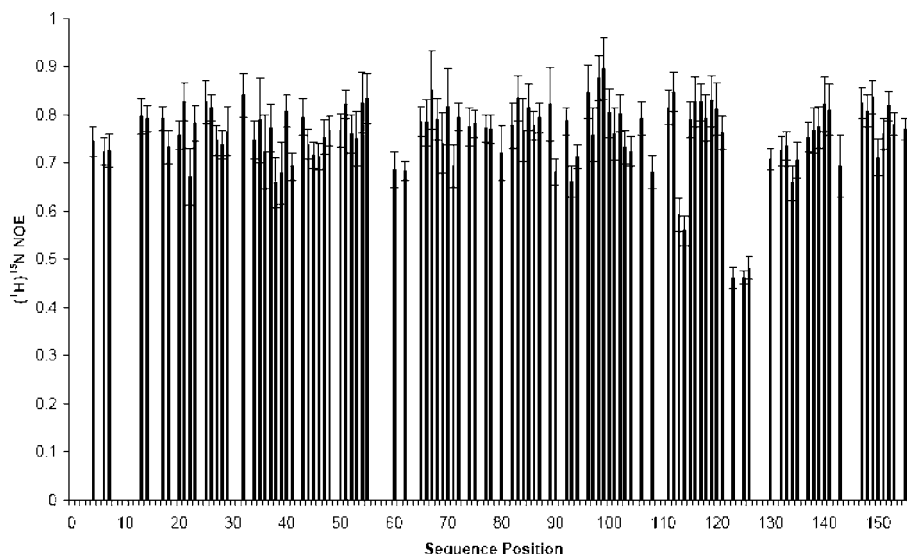


Figure 4 $\{^1\text{H}\}-^{15}\text{N}$ heteronuclear NOE values for Gly m 4

Missing values could not be determined because of signal overlap; errors were calculated as described in [17]. The increased flexibility of the N-terminus of strand $\beta 7$ and of the connecting loop $\beta 7 \rightarrow \alpha 3$ is clearly visible.

m 4 and Bet v 1 in this region is not as high as that in the P-loop region. Due to the shift of helix $\alpha 3$ along its axis, the distances of the amino acids located in helices $\alpha 1$ and $\alpha 2$ to those in $\alpha 3$ differ slightly from the corresponding distances in Bet v 1, but these differences do not impair the general similarity of the protein surfaces. Helix $\alpha 1$ was identified as part of an IgE epitope of Bet v 1 by mimotope selection [2], and Asn²⁸ located in helix $\alpha 2$ was shown to be crucial for IgE binding to Pru av 1 [35]. These findings strongly support the notion that the conserved surface patch represents another cross-reactive epitope on Gly m 4.

The third conserved surface patch consists of residues located in strands $\beta 1$, $\beta 7$, $\beta 6$, $\beta 5$ and adjacent loops $\beta 5-\beta 6/\beta 7-\alpha 3$. It covers a coherent area of $\sim 1100 \text{ \AA}^2$. The spatial arrangement of these β -strands is highly similar in Bet v 1 and Gly m 4. This putative cross-reactive epitope is probably the one showing the highest flexibility, as it contains a number of residues located in the extended loop $\beta 7-\alpha 3$. The respective loop in Bet v 1 contains Asp¹²⁵, which seems to play a role in IgE binding, as its mutation modulates IgE reactivity [3].

These surface areas largely overlap with the three surface patches conserved among Bet v 1 and other tree pollen allergens from the order *fagales* like Aln g 1, Cor a 1 or Car b 1. These potential cross-reactive epitopes [7] are, likely, not only responsible for the pollen-food allergy syndrome, but also for cross-reactivity among tree pollen.

In addition to these three highly coherent epitopes, a fourth area of conserved residues covering $\sim 1050 \text{ \AA}^2$ is found in Gly m 4. It includes residues located in strands $\beta 1$, $\beta 7$ and $\beta 6$ as well as in loops $\beta 4-\beta 5$ and $\beta 6-\beta 7$. Orientation of the β -strands and the conformation of the loops are identical in Gly m 4 and Bet v 1. Two sequence stretches in Bet v 1 identified by mimotope selection both include residues from this area [36], and a six-

point mutant of Bet v 1 including mutation of Ser¹¹², which corresponds to the N-terminal serine residue of $\beta 7$ in Gly m 4, showed drastically reduced IgE binding [3]. S112C single mutants of Bet v 1 and Pru av 1 and a P108G mutant of Bet v 1 also showed reduced IgE reactivity [31,33].

From clinical studies, Gly m 4 [11] as well as Api g 1 [37] causes more severe symptoms than, for example, Pru av 1 [38] or Cor a 1 [39]. As the unfolding temperatures of Bet v 1, Gly m 4 and Pru av 1 are all in the range of $65-75^\circ\text{C}$ as determined by CD spectroscopy (H. Berkner, unpublished work), differential stability is highly likely not the major cause of the differences in the severity of symptoms.

Conclusion

Cross-reactive epitopes on homologous allergens result from a highly similar tertiary structure in combination with local or extended sequence similarity. We show here that, apart from a few slight conformational differences, the three-dimensional structure of Gly m 4 is virtually identical with that of the major birch pollen allergen Bet v 1. Generally, our results also confirm that the grade of total sequence similarity correlates with the grade of cross-reactivity, because Bet v 1, Mal d 1 and Pru av 1 with sequence similarities of 66–69% to Gly m 4 show significant cross-reactivity, whereas Api g 1 and Dau c 1 with sequence similarity of 58 and 56% to Gly m 4 did not exhibit any cross-reactivity with the soya-bean allergen. Sequence similarity, however, is not able to predict cross-reactivity in detail: Bet v 1, the protein with the lowest sequence similarity to Gly m 4 in the first group, shows the highest cross-reactivity, whereas the protein with the highest sequence similarity, Pru av 1, even fails to inhibit Gly m 4 in some of the studied patients.

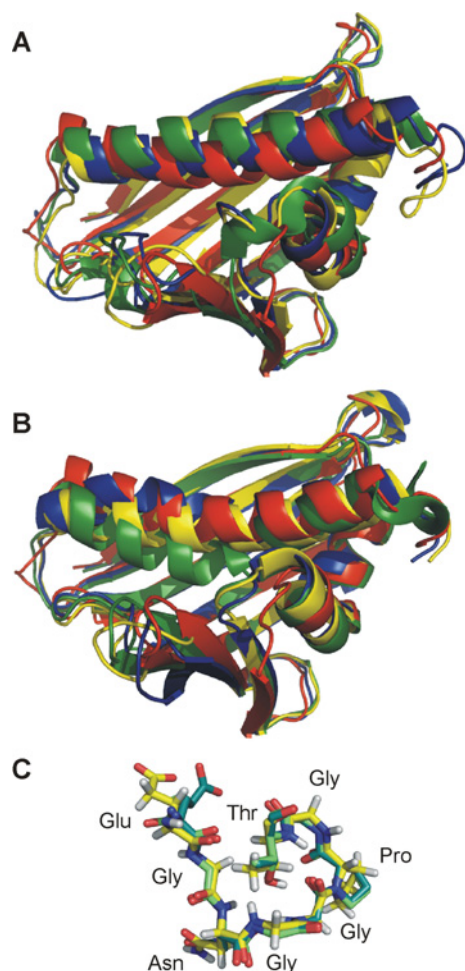


Figure 5 ?????

Overlay of the lowest energy solution structure of Gly m 4 (2K7H, red) with (A) the structures of the major birch pollen allergen Bet v 1 (1BV1, blue), the major cherry allergen Pru av 1 (1E09, yellow) and the celery allergen Api g 1 (2BK0, green) and (B) the structures of three Bet v 1 homologous proteins from yellow lupine, namely LIPR10.1A (1ICX, blue), LIPR10.1B (1IFV, yellow) and LIPR10.2A (1XDF, green). (C) Overlay of the P-loop regions of Gly m 4 (yellow), Bet v 1 (blue-grey) and LIPR10.1B (light green); oxygen atoms are marked in red and nitrogen atoms in blue. While the P-loop regions of Gly m 4, Bet v 1 and LIPR10.1B are virtually identical, the agreement of the structures of Gly m 4 and the lupine proteins is clearly better than the agreement of the structures of Gly m 4 and the allergens Bet v 1, Pru av 1 and Api g 1, mainly due to the bent C-terminal helix.

In the context of our findings that many of the structural features of Gly m 4 are particularly similar to the PR10 proteins of yellow lupine, it would be interesting to investigate general allergenicity and cross-reactivity of Gly m 4 with these highly homologous proteins (sequence similarities of 74–84%). In fact, white lupine has been identified as a food allergen source [40]. Immunoblotting as well as an *in silico* approach [41,42] supports the notion that a PR10 protein of white lupine with sequence similarities of 75–94% to the yellow lupine proteins is responsible for lupine allergy.

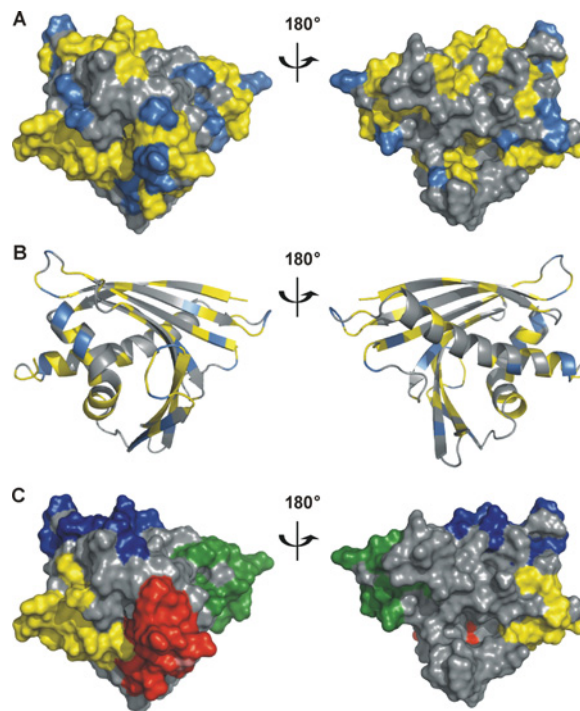


Figure 6 ?????

(A) Surface representation of the lowest energy structure of Gly m 4 with residues conserved in Bet v 1 shown in blue and residues conserved in Bet v 1, Mal d 1 and Pru av 1 shown in yellow. (B) Cartoon representation of Gly m 4 in the same orientation and with the same labelling as in (A). (C) Surface representation of the lowest energy structure of Gly m 4 in the same orientation as in (A) with putative cross-reactive epitopes marked in red (P-loop), yellow (helices $\alpha 1$, $\alpha 2$ and $\alpha 3$), green (strands $\beta 1$, $\beta 5$, $\beta 6$ and $\beta 7$) and blue (strands $\beta 1$, $\beta 6$ and $\beta 7$).

Our results also demonstrate the importance of considering patient-specificity of the IgE reaction for the design of hypo-allergenic proteins for immunotherapy. The inhibitory potential of Pru av 1 depends on the specific patients tested, suggesting that these patients recognize different epitopes [2]. Further efforts for a detailed characterization of the conformational IgE epitopes of relevant allergens are thus necessary to understand the clinical complexity of the phenomenon at the molecular level and to pave the way for a safe and patient-tailored immunotherapy.

ACKNOWLEDGEMENTS

We thank Nadine Herz and Ramona Heissmann for expert technical assistance and Katrin Lehmann for her contributions in the early stages of this work.

FUNDING

This work was supported by the Deutsche Forschungsgemeinschaft [grant number RO 617/11-4 (to P. R.)], by a doctoral fellowship awarded to P. N. from the Fonds des Verbandes der Chemischen

Industrie in co-operation with the Bundesministerium für Bildung und Forschung [grant number ???], and by the Elitenetwork of Bavaria [grant number ???].

REFERENCES

- Vieths, S., Scheurer, S. and Ballmer-Weber, B. (2002) Current understanding of cross-reactivity of food allergens and pollen. *Ann. N.Y. Acad. Sci.* 964, 47–68
- Mittag, D., Batori, V., Neudecker, P., Wiche, R., Friis, E. P., Ballmer-Weber, B. K., Vieths, S. and Roggen, E. L. (2006) A novel approach for investigation of specific and cross-reactive IgE epitopes on bet v 1 and homologous food allergens in individual patients. *Mol. Immunol.* 43, 268–278
- Ferreira, F., Ebner, C., Kramer, B., Casari, G., Briza, P., Kungl, A. J., Grimm, R., Jahn-Schmid, B., Breiteneder, H., Kraft, D. et al. (1998) Modulation of IgE reactivity of allergens by site-directed mutagenesis: potential use of hypoallergenic variants for immunotherapy. *FASEB J.* 12, 231–242
- Murzin, A. G., Brenner, S. E., Hubbard, T. and Chothia, C. (1995) SCOP: a structural classification of proteins database for the investigation of sequences and structures. *J. Mol. Biol.* 247, 536–540
- Fernandes, H., Pasternak, O., Bujacz, G., Bujacz, A., Sikorski, M. M. and Jaskolski, M. (2008) Lupinus luteus pathogenesis-related protein as a reservoir for cytokinin. *J. Mol. Biol.* 378, 1040–1051
- Berkner, H., Schweimer, K., Matecko, I. and Rosch, P. (2008) Conformation, catalytic site, and enzymatic mechanism of the PR10 allergen-related enzyme norcoclaurine synthase. *Biochem. J.* 413, 281–290
- Gajhede, M., Osmark, P., Poulsen, F. M., Ipsen, H., Larsen, J. N., Joost van Neerven, R. J., Schou, C., Lowenstein, H. and Spangfort, M. D. (1996) X-ray and NMR structure of bet v 1, the origin of birch pollen allergy. *Nat. Struct. Biol.* 3, 1040–1045
- Neudecker, P., Schweimer, K., Nerkamp, J., Scheurer, S., Vieths, S., Sticht, H. and Rosch, P. (2001) Allergic cross-reactivity made visible: solution structure of the major cherry allergen pru av 1. *J. Biol. Chem.* 276, 22756–22763
- Schirmer, T., Hoffmann-Sommergrube, K., Susani, M., Breiteneder, H. and Markovic-Housley, Z. (2005) Crystal structure of the major celery allergen api g 1: molecular analysis of cross-reactivity. *J. Mol. Biol.* 351, 1101–1109
- Crowell, D. N., John, M. E., Russell, D. and Amasino, R. M. (1992) Characterization of a stress-induced, developmentally regulated gene family from soybean. *Plant Mol. Biol.* 18, 459–466
- Kleine-Tebbe, J., Vogel, L., Crowell, D. N., Haustein, U. F. and Vieths, S. (2002) Severe oral allergy syndrome and anaphylactic reactions caused by a bet v 1-related PR-10 protein in soybean, SAM22. *J. Allergy Clin. Immunol.* 110, 797–804
- Mittag, D., Vieths, S., Vogel, L., Becker, W. M., Rihs, H. P., Helbling, A., Wuthrich, B. and Ballmer-Weber, B. K. (2004) Soybean allergy in patients allergic to birch pollen: clinical investigation and molecular characterization of allergens. *J. Allergy Clin. Immunol.* 113, 148–154
- King, T. P., Hoffman, D., Lowenstein, H., Marsh, D. G., Platts-Mills, T. A. and Thomas, W. (1995) Allergen nomenclature. *Allergy* 50, 765–774
- Ballmer-Weber, B. K. and Vieths, S. (2008) Soy allergy in perspective. *Curr. Opin. Allergy Clin. Immunol.* 8, 270–275
- Neudecker, P., Lehmann, K. and Rosch, P. (2003) Sequence-specific ^1H , ^{13}C and ^{15}N resonance assignments of SAM22, an allergenic stress-induced protein from soy bean. *J. Biomol. NMR* 26, 191–192
- Cavanagh, J., Fairbrother, W. J., Palmer, III, A. G. and Skelton, N. J. (2006) *Protein NMR Spectroscopy: Principles and Practice*. Academic Press, San Diego, CA
- Farrow, N. A., Muhandiram, R., Singer, A. U., Pascal, S. M., Kay, C. M., Gish, G., Shoelson, S. E., Pawson, T., Forman-Kay, J. D. and Kay, L. E. (1994) Backbone dynamics of a free and phosphopeptide-complexed Src homology 2 domain studied by ^{15}N NMR relaxation. *Biochemistry* 33, 5984–6003
- Johnson, B. A. and Blevins, R. A. (1994) NMRView: a computer program for the visualization and analysis of NMR data. *J. Biomol. NMR* 4, 603–614
- Cornilescu, G., Delaglio, F. and Bax, A. (1999) Protein backbone angle restraints from searching a database for chemical shift and sequence homology. *J. Biomol. NMR* 13, 289–302
- Tjandra, N., Omichinski, J. G., Gronenborn, A. M., Clore, G. M. and Bax, A. (1997) Use of dipolar ^1H - ^{15}N and ^1H - ^{13}C couplings in the structure determination of magnetically oriented macromolecules in solution. *Nat. Struct. Biol.* 4, 732–738
- Laskowski, R. A., Rullmann, J. A., MacArthur, M. W., Kaptein, R. and Thornton, J. M. (1996) AQUA and PROCHECK-NMR: programs for checking the quality of protein structures solved by NMR. *J. Biomol. NMR* 8, 477–486
- Koradi, R., Billeter, M. and Wüthrich, K. (1996) MOLMOL: a program for display and analysis of macromolecular structures. *J. Mol. Graphics* 14, 51–55
- Hubbard, S. J. and Thornton, J. M. (1993) NACCESS
- Ballmer-Weber, B. K., Holzhauser, T., Scibilia, J., Mittag, D., Zisa, G., Ortolani, C., Oesterballe, M., Poulsen, L. K., Vieths, S. and Bindslev-Jensen, C. (2007) Clinical characteristics of soybean allergy in Europe: a double-blind, placebo-controlled food challenge study. *J. Allergy Clin. Immunol.* 119, 1489–1496
- Laemmli, U. K. (1970) Cleavage of structural proteins during the assembly of the head of bacteriophage T4. *Nature* 227, 680–685
- Neudecker, P., Lehmann, K., Nerkamp, J., Haase, T., Wangorsch, A., Fotisch, K., Hoffmann, S., Rosch, P., Vieths, S. and Scheurer, S. (2003) Mutational epitope analysis of pru av 1 and api g 1, the major allergens of cherry (*Prunus avium*) and celery (*Apium graveolens*): correlating IgE reactivity with three-dimensional structure. *Biochem. J.* 376, 97–107
- Wangorsch, A., Ballmer-Weber, B. K., Rosch, P., Holzhauser, T. and Vieths, S. (2007) Mutational epitope analysis and cross-reactivity of two isoforms of api g 1, the major celery allergen. *Mol. Immunol.* 44, 2518–2527
- Biesiadka, J., Bujacz, G., Sikorski, M. M. and Jaskolski, M. (2002) Crystal structures of two homologous pathogenesis-related proteins from yellow lupine. *J. Mol. Biol.* 319, 1223–1234
- Pasternak, O., Biesiadka, J., Dolot, R., Handschuh, L., Bujacz, G., Sikorski, M. M. and Jaskolski, M. (2005) Structure of a yellow lupin pathogenesis-related PR-10 protein belonging to a novel subclass. *Acta Crystallogr. Sect. D Biol. Crystallogr.* 61, 99–107
- Markovic-Housley, Z., Degano, M., Lamba, D., von Roepenack-Lahaye, E., Clemens, S., Susani, M., Ferreira, F., Scheiner, O. and Breiteneder, H. (2003) Crystal structure of a hypoallergenic isoform of the major birch pollen allergen bet v 1 and its likely biological function as a plant steroid carrier. *J. Mol. Biol.* 325, 123–133
- Spangfort, M. D., Mirza, O., Ipsen, H., Joost van Neerven, R. J., Gajhede, M. and Larsen, J. N. (2003) Dominating IgE-binding epitope of bet v 1, the major allergen of birch pollen, characterized by X-ray crystallography and site-directed mutagenesis. *J. Immunol.* 171, 3084–3090
- Neudecker, P., Lehmann, K., Nerkamp, J., Haase, T., Wangorsch, A., Fotisch, K., Hoffmann, S., Rosch, P., Vieths, S. and Scheurer, S. (2003) Mutational epitope analysis of pru av 1 and api g 1, the major allergens of cherry (*Prunus avium*) and celery (*Apium graveolens*): correlating IgE reactivity with three-dimensional structure. *Biochem. J.* 376, 97–107
- Scheurer, S., Son, D. Y., Boehm, M., Karamloo, F., Franke, S., Hoffmann, A., Haustein, D. and Vieths, S. (1999) Cross-reactivity and epitope analysis of pru a 1, the major cherry allergen. *Mol. Immunol.* 36, 155–167

Q4



- 34 Mirza, O., Henriksen, A., Ipsen, H., Larsen, J. N., Wissenbach, M., Spangfort, M. D. and Gajhede, M. (2000) Dominant epitopes and allergic cross-reactivity: complex formation between a Fab fragment of a monoclonal murine IgG antibody and the major allergen from birch pollen bet v 1. *J. Immunol.* 165, 331–338
- 35 Wiche, R., Gubesch, M., Konig, H., Fotisch, K., Hoffmann, A., Wangorsch, A., Scheurer, S. and Vieths, S. (2005) Molecular basis of pollen-related food allergy: identification of a second cross-reactive IgE epitope on Pru av 1, the major cherry (*Prunus avium*) allergen. *Biochem. J.* 385, 319–327
- 36 Ganglberger, E., Grunberger, K., Sponer, B., Radauer, C., Breiteneder, H., Boltz-Nitulescu, G., Scheiner, O. and Jensen-Jarolim, E. (2000) Allergen mimotopes for 3-dimensional epitope search and induction of antibodies inhibiting human IgE. *FASEB J.* 14, 2177–2184
- 37 Luttkopf, D., Ballmer-Weber, B. K., Wuthrich, B. and Vieths, S. (2000) Celery allergens in patients with positive double-blind placebo-controlled food challenge. *J. Allergy Clin. Immunol.* 106, 390–399
- 38 Reuter, A., Lidholm, J., Andersson, K., Ostling, J., Lundberg, M., Scheurer, S., Enrique, E., Cistero-Bahima, A., San Miguel-Moncin, M., Ballmer-Weber, B. K. and Vieths, S. (2006) A critical assessment of allergen component-based *in vitro* diagnosis in cherry allergy across Europe. *Clin. Exp. Allergy* 36, 815–823
- 39 Luttkopf, D., Muller, U., Skov, P. S., Ballmer-Weber, B. K., Wuthrich, B., Skamstrup Hansen, K., Poulsen, L. K., Kastner, M., Hausteiner, D. and Vieths, S. (2002) Comparison of four variants of a major allergen in hazelnut (*Corylus avellana*) Cor a 1.04 with the major hazel pollen allergen Cor a 1.01. *Mol. Immunol.* 38, 515–525
- 40 Peeters, K. A., Nordlee, J. A., Penninks, A. H., Chen, L., Goodman, R. E., Bruijnzeel-Koomen, C. A., Hefle, S. L., Taylor, S. L. and Knulst, A. C. (2007) Lupine allergy: not simply cross-reactivity with peanut or soy. *J. Allergy Clin. Immunol.* 120, 647–653
- 41 Faeste, C. K., Lovik, M., Wiker, H. G. and Egeas, E. (2004) A case of peanut cross-allergy to lupine flour in a hot dog bread. *Int. Arch. Allergy Immunol.* 135, 36–39
- 42 Guarneri, F., Guarneri, C. and Benveniste, S. (2005) Identification of potentially cross-reactive peanut-lupine proteins by computer-assisted search for amino acid sequence homology. *Int. Arch. Allergy Immunol.* 138, 273–277
- 43 Chenna, R., Sugawara, H., Koike, T., Lopez, R., Gibson, T. J., Higgins, D. G. and Thompson, J. D. (2003) Multiple sequence alignment with the Clustal series of programs. *Nucleic Acids Res.* 31, 3497–3500

Received 21 August 2008/1 October 2008; accepted 3 October 2009

Published as Immediate Publication 3 October 2008, doi 10.1042/BSR20080117

Einzelarbeit B

Hanna Berkner, Julia Engelhorn, David K. Liscombe, Kristian Schweimer, Birgitta M. Wöhl, Peter J. Facchini, Paul Rösch und Irena Matečko (2007): High-yield expression and purification of isotopically labeled norcoclaurine synthase, a Bet v 1-homologous enzyme, from *Thalictrum flavum* for NMR studies. *Protein Expression and Purification* **56**(2), 197-204

High-yield expression and purification of isotopically labeled norcoclaurine synthase, a Bet v 1-homologous enzyme, from *Thalictrum flavum* for NMR studies

Hanna Berkner^a, Julia Engelhorn^a, David K. Liscombe^b, Kristian Schweimer^a,
Birgitta M. Wöhrl^a, Peter J. Facchini^b, Paul Rösch^a, Irena Matečko^{a,*}

^a Department of Biopolymers and Research Center for Bio-Macromolecules, Universität Bayreuth, Universitätsstraße 30, 95440 Bayreuth, Germany

^b Department of Biological Sciences, University of Calgary, Calgary, Alta. T2N 1N4, Canada

Received 12 June 2007, and in revised form 16 July 2007

Available online 24 August 2007

Abstract

The enzyme norcoclaurine synthase (NCS) found in the common meadow rue, *Thalictrum flavum*, and other plants shows sequence homology to members of the class 10 of pathogenesis related (PR 10) proteins that contains allergens such as the major birch pollen allergen Bet v 1, the major cherry allergen Pru av 1, and the major apple allergen Mal d 1. The enzyme is involved in the plant's secondary metabolism and is required for the production of bioactive secondary metabolites like morphine. Whereas the physiological function of PR 10 class allergens is still unknown, NCS activity has been studied in detail. Investigation of the structural properties of NCS by NMR spectroscopy can thus not only provide new information concerning the reaction mechanism of the enzyme, but is also expected to help clarify the long standing and heavily debated question on the physiological function as well as the reasons for the allergenic potential of members of this protein family. As the first important step towards the three-dimensional solution structure, we optimized expression of recombinant NCS in *Escherichia coli* and established an efficient purification protocol yielding high amounts of pure isotopically labeled active enzyme. The identity of NCS was confirmed by electrospray ionization mass spectrometry, and activity of the purified enzyme was determined by an assay detecting the radiolabeled reaction product. Spectroscopic analysis by NMR spectroscopy showed that the protein was properly folded with well defined tertiary structure.

© 2007 Elsevier Inc. All rights reserved.

Keywords: Norcoclaurine synthase; PR-10 protein family; Protein purification; Isotope labeling; NMR

The enzyme norcoclaurine synthase (NCS)¹ found among other plants in opium poppy (*Papaver somniferum*) and in the common meadow rue (*Thalictrum flavum*) shows up to 38% sequence identity with members of class 10 of pathogenesis related (PR 10) proteins

which includes the family of Bet v 1 homologous allergens [1–3]. PR proteins represent a group of plant proteins divided into 14 classes [4]. Their main common characteristic is the induction of their expression under stress conditions [5].

* Corresponding author. Fax: +49 921 553544.

E-mail address: irena.matecko@uni-bayreuth.de (I. Matečko).

¹ Abbreviations used: NCS, norcoclaurine synthase; NMR, nuclear magnetic resonance; PR 10, class 10 of pathogenesis related proteins; 4-HPAA, 4-hydroxyphenylacetaldehyde; PCR, polymerase chain reaction; His₆-tag, hexahistidine-tag; LB medium, Luria–Bertani medium; SDS, sodium dodecyl sulfate; PAGE, polyacrylamide gel electrophoresis; OD₆₀₀, optical density at 600 nm; IPTG, isopropyl-β-D-thiogalactoside; IMAC, immobilized metal ion affinity chromatography; FPLC, fast performance liquid chromatography; ESI-MS, electrospray ionization mass spectrometry; TLC, thin-layer chromatography; DTT, dithiothreitol; HSQC, heteronuclear single quantum coherence; MALDI-TOF-MS, matrix assisted laser desorption ionization time-of-flight mass spectrometry.

So far, little is known about the physiological function of members of the Bet v 1 family in plants, but a growing number of people, at present about 25% of the total population of the industrialized countries [6–8], experiences the effect these proteins can have on the human immune system: they cause type I allergies involving weaker symptoms like allergic rhinitis as well as asthma and sometimes even severe anaphylactic reactions. The molecular mechanism leading to these symptoms has been investigated in detail. The dimerization of allergen-specific type E immunoglobulins provoked by binding of two allergen epitopes represents the crucial step in a signal cascade that finally leads to the release of mediators like histamine and serotonin [9–12]. A promising approach to develop a therapy fighting this type of allergy is the hyposensibilisation of patients using so-called hypoallergens. These artificially designed proteins have the ability to provoke an immune response involving type G instead of type E immunoglobulins and thus to prevent the subsequent signal cascade and the resulting symptoms [13]. An essential prerequisite for the construction of such a hypoallergen is detailed information about the structure of the naturally occurring allergen.

Three-dimensional structures of a number of allergens are already known, among them the structure of the major birch pollen allergen Bet v 1 [14,15], the cherry allergen Pru av 1 [16], the minor birch pollen allergen Bet v 4 [17], and others [18]. Based on their amino acid sequence homology to Bet v 1, several other pollen allergens and also different food allergens like Pru av 1 or Mal d 1 are grouped into the family of Bet v 1-homologous allergens. The members of this family share a number of common structural features. Consisting of 153–160 amino acids, their sequence contains a highly conserved P-loop motive with a G-X-G-G-X-G-T consensus sequence [19,20]. Whereas the P-loop is also present in NCS, this enzyme comprising 210 amino acids contains extensions at both the amino- and the carboxy-terminus as compared to the members of the Bet v 1 family. The determination of the three-dimensional structure of NCS which is unique in the Bet v 1 family as its function in the plant is known can thus shed some light on the evolutionary connections between enzymatically active members of the PR 10 class and proteins such as Bet v 1 and Pru av 1 for which an enzymatic activity has not been established yet. Such studies are also intended to help answer the therapeutically relevant question about the structural basis of allergic reactions.

Additionally, structural data can also provide new information about the reaction mechanism of the enzyme. NCS catalyzes the first step in benzylisoquinoline alkaloid synthesis which consists of the condensation of dopamine and 4-hydroxyphenylacetaldehyde (4-HPAA) to norcoclaurine [21]. Benzylisoquinoline alkaloids form an interesting group of plant secondary metabolites because many of them show pharmacological activity, for example the analgesics morphine and codeine, the muscle relaxant papaverine, and the antibiotic sanguinarine [21]. The common meadow rue, *T. flavum*, produces the antimicrobial berber-

ine and magnoflorine which has been suggested to possess anti-HIV activity [22,23]. In 2004, the gene coding for NCS was isolated from a *T. flavum* cDNA library, the gene was cloned and the protein was expressed in *Escherichia coli*. Enzyme activity assays were carried out using the crude extract obtained after cell lysis [1]. With the aim to determine the NMR solution structure of this in many respects interesting protein, we optimized the expression of NCS in *E. coli* and established a purification protocol yielding sufficient amounts of properly folded and highly pure isotopically labeled enzyme.

Materials and methods

Plasmid construction

The plasmid pET29b- Δ 19NCS [1] containing the cDNA of *T. flavum* norcoclaurine synthase coding for amino acids 20–210 cloned with *Bam*HI/*Xho*I was used as a template for construction of pET29b- Δ 29NCS coding for amino acids 30–210. The corresponding gene fragment was amplified by PCR using a forward primer including an *Nde*I site (underlined; 5'-AGGTAGGCATATGCTGCACCACCA GGCATAATAAACC-3') and a reverse primer containing an *Xho*I site (underlined; 3'-GTGGTGCTCGAGGACTGTTATTATTGC-5'). The temperature cycling for 35 cycles was set up as follows: 94 °C for 30 s, 54 °C for 30 s and 72 °C for 1 min. The amplified fragment was inserted into expression vector pET29b (Novagen, Madison, USA) using *Nde*I/*Xho*I restriction sites, and successful cloning was confirmed by DNA sequencing. In contrast to plasmid pET29b- Δ 19NCS coding for an amino-terminal S-tag fusion with a carboxy-terminal hexahistidine-tag (His₆-tag), expression of pET29b- Δ 29NCS yields only C-terminally His₆-tagged Δ 29NCS.

Expression in LB medium

Plasmid pET29b- Δ 19NCS was either expressed in *E. coli* ER2566pLysS [1] or transformed into *E. coli* BL21(DE3) and *E. coli* Rosetta(DE3) host cells. For expression of Δ 29NCS, plasmid pET29b- Δ 29NCS was transformed into *E. coli* Rosetta(DE3) host cells. After transformation host cells were grown on LB agar containing the appropriate antibiotics (ER2566pLysS/pET29b- Δ 19NCS and BL21(DE3)/pET29b- Δ 19NCS: 20 µg/ml kanamycin; Rosetta(DE3)/pET29b- Δ 29NCS: 20 µg/ml kanamycin, 34 µg/ml chloramphenicol) for selection at 37 °C overnight. For induction tests, 10–20 mL LB medium containing the respective antibiotics was inoculated with the expression strain and grown overnight at 37 °C. This starter culture was used to inoculate 100 mL of the same medium to an OD₆₀₀ of 0.2. Cells were then incubated at 37 °C under continuous shaking. At an OD₆₀₀ of 0.6, expression was induced by addition of 1 mM IPTG followed by continuous incubation at 37 °C for 4–6 h. To monitor the expression, aliquots containing equal amounts of cells were

taken from the culture every hour and applied to a 15% SDS–polyacrylamide gel. For preparative expression, the same procedure was carried out using final culture volumes up to 2.5 L. Cells were harvested 4–6 h after induction by centrifugation at 5000g for 15 min at 4 °C and the cell pellet was stored at –20 °C.

Preparative expression in M9 minimal medium

For expression of $\Delta 19\text{NCS}$ and $\Delta 29\text{NCS}$ in minimal medium, 100 mL of M9 [24] containing 2 mM MgSO_4 , 10 μM Fe(III) citrate , 0.1 mM CaCl_2 , 4 g/L glucose, trace element solution TS2 [25], MEM vitamin solution (Gibco, Invitrogen, Karlsruhe), 20 $\mu\text{g/mL}$ kanamycin, 34 $\mu\text{g/mL}$ chloramphenicol and 1.5 g/L $(\text{NH}_4)_2\text{SO}_4$ as the only source of nitrogen were inoculated with *E. coli* Rosetta(*DE3*)/pET29b- $\Delta 29\text{NCS}$ and grown overnight at 37 °C. Two liter of the same medium were inoculated with the overnight culture to an OD_{600} of 0.2 and incubated under shaking at 37 °C. Induction was carried out at an OD_{600} of 0.6 by adding 1 mM IPTG and cells were harvested after 5–6 h of expression and stored as described above. Expression of ^{15}N -labeled protein was achieved replacing $(\text{NH}_4)_2\text{SO}_4$ by $(^{15}\text{NH}_4)_2\text{SO}_4$.

Purification of $\Delta 19\text{NCS}$ and $\Delta 29\text{NCS}$ by affinity chromatography

Frozen cells were resuspended in binding buffer (20 mM sodium phosphate buffer pH 7.8, 500 mM NaCl, 20 mM imidazole; $\Delta 19\text{NCS}$: 5 mL/g pellet; $\Delta 29\text{NCS}$: 7 mL/g pellet) containing lysozyme, DNase, and one protease inhibitor cocktail tablet (Complete, EDTA-free, Roche) and stirred on ice for 30 min followed by sonication on ice. The lysate was centrifuged at 19,100g for 45 min at 4 °C. The pellet was resuspended in binding buffer, stirred on ice, sonicated, and centrifuged again as described above, and supernatants were pooled. Purification of the soluble his-tagged protein was performed by immobilized metal ion affinity chromatography (IMAC) on an ÄKTA purifier 10-FPLC system (Amersham Biosciences, Freiburg, Germany). The supernatant was loaded on a 5 mL HisTrap column (Amersham Biosciences, Freiburg, Germany) pre-equilibrated with binding buffer. After a washing step with binding buffer, the bound protein was eluted using a stepwise gradient including 5, 10, 15, 20, 25, 30, 50, and 100% elution buffer (20 mM sodium phosphate buffer pH 7.8, 500 mM NaCl, 300 mM imidazole). The collected 5 mL fractions were analyzed by SDS–PAGE. In case of $\Delta 29\text{NCS}$ the fractions containing the protein were dialyzed as first step against 5 L 20 mM sodium phosphate buffer pH 7.8 containing 100 mM NaCl and then twice against 5 L Millipore water, every step at least for 3 h at 4 °C. After shock-freezing in liquid nitrogen, the protein was lyophilized and stored at 4 °C. In case of $\Delta 19\text{NCS}$, the S-tag-fusion was subjected to thrombin cleavage.

Thrombin cleavage of $\Delta 19\text{NCS}$

After affinity chromatography, the fractions containing $\Delta 19\text{NCS}$ were pooled and dialyzed against 5 L of binding buffer without imidazole for at least 3 h at 4 °C. Thrombin cleavage removing the amino-terminal S-tag from the $\Delta 19\text{NCS}$ -fusion was carried out at room temperature overnight during dialysis against 5 L fresh binding buffer without imidazole by adding 0.2 U thrombin (Novagen, Madison, USA) per mg protein inside the dialysis membrane (Spectra/Por, MWCO 3500, ROTH, Karlsruhe, Germany). Successful cleavage was confirmed by SDS–PAGE. To remove the cleaved S-tag, the sample was again subjected to nickel ion affinity chromatography under the conditions described above. Subsequent dialysis and lyophilization was also carried out as described above.

Gel electrophoresis

Denaturing SDS–PAGE (15% w/v polyacrylamide) was performed as described by Lämmli [26]. Samples were denatured by dilution with Roti-Load 1 (ROTH, Karlsruhe, Germany) and incubation at 95 °C for 5–10 min. Proteins were visualized by Coomassie blue staining.

Protein mass fingerprint and N-terminal sequencing

In order to identify the fragments of $\Delta 19\text{NCS}$ visible on the SDS–polyacrylamide gel before and after thrombin cleavage, gel pieces with 1 mm diameter were cut from the corresponding band and subjected to protein mass fingerprinting including tryptic digestion, determination of the resulting peptide masses by MALDI-TOF-MS and analysis of the covered protein sequence using the MASCOT program. N-terminal sequencing of the $\Delta 19\text{NCS}$ degradation product was carried out by Edman degradation on an ABI Procise 491-Sequencer. Both methods were performed by ZMMK Koeln, Zentrale Bioanalytik.

Electrospray ionization mass spectrometry

Mass spectra were recorded on a LCT Micromass time-of-flight electrospray mass spectrometer from Waters with an electrospray interface, in positive ion-mode. The spray tip potential was 2000 V, sample cone potential was 50 V, and desolvation temperature 150 °C. The software used for recording and processing the data was Masslynx v3.4 (Waters). For data analysis MagTran 1.02 programme was used [27].

Stock solution was prepared by dissolving lyophilized protein in 1 mM ammonium acetate buffer pH 7.0. Concentration of the final stock solution was 1 mM. For the ESI-MS experiments 100 μM protein solution in 1 mM ammonium acetate, 5% ethanol and 0.1% acetic acid was used.

Enzyme activity assay

Enzymatic activity of purified $\Delta 29\text{NCS}$ was tested as described [1]. Lyophilized protein was resuspended in assay buffer (100 mM Tris-HCl, pH 7.0; 12 mM β -mercaptoethanol). Protein concentration was determined according to the Bradford method [28]. Each assay contained 15 μg of pure $\Delta 29\text{NCS}$, 311 pmol (187 Bq) [$8\text{-}^{14}\text{C}$] dopamine and 25 nmol 4-HPAA in a total volume of 30 μl assay buffer. The negative control contained protein solution boiled for 10 min prior to assay setup. The assays were incubated for 1.5 h at 37 °C. Subsequently, the whole assay mixture was subjected to thin-layer chromatography (TLC) and radiolabeled compounds were visualized by autoradiography.

NMR spectroscopy

NMR samples were prepared by dissolving the required amount of lyophilized ^{15}N -labeled $\Delta 29\text{NCS}$ in 20 mM sodium phosphate buffer pH 7.0, 2.5 mM DTT, 0.04% NaN_3 and 10% D_2O . Final sample concentration was determined by measuring absorption of the protein solution at 280 nm. Two-dimensional ^{15}N - ^1H HSQC NMR spectra were recorded according to the FHSQC-scheme [29] on a Bruker Avance 800 MHz NMR spectrometer equipped with a triple-resonance cryoprobe with pulsed field-gradient capabilities.

Results and discussion

Optimization of $\Delta 19\text{NCS}$ expression in *E. coli*

Three different *E. coli* host strains were tested for expression of $\Delta 19\text{NCS}$ (25.8 kDa) comprising an amino-terminal S-tag followed by a thrombin cleavage site and a carboxy-terminal His₆-tag. *E. coli* ER2566pLysS was already used as host strain for $\Delta 19\text{NCS}$ expression [1] in order to investigate the enzymatic activity of recombinant NCS present in the cell lysate. Whereas only small amounts of enzyme are necessary to test its activity, highly concentrated and

pure protein is needed for structural studies by NMR spectroscopy. For this reason, expression of $\Delta 19\text{NCS}$ using plasmid pET29b- $\Delta 19\text{NCS}$ had to be optimized. In addition to *E. coli* ER2566pLysS, *E. coli* BL21(DE3) and *E. coli* Rosetta(DE3) were tested as alternative expression hosts (Fig. 1). Only low expression rates and, as consequence, no accumulation of significant amounts of protein over expression time could be observed for *E. coli* ER2566pLysS (Fig. 1a) and *E. coli* BL21(DE3) (Fig. 1b). However, expression in *E. coli* Rosetta(DE3) yielded high amounts of $\Delta 19\text{NCS}$ augmenting over expression time (Fig. 1c). As the sequence coding for $\Delta 19\text{NCS}$ is identical to that of the NCS gene from *T. flavum*, differences in codon usage between the originating organism and the expression host might lead to problems during translation of the recombinant protein due to the lack of certain tRNAs [30]. Analyzing the sequence coding for $\Delta 19\text{NCS}$ including both tags, a number of rare codons can be identified, especially coding for leucine and isoleucine. The special feature of *E. coli* Rosetta(DE3) is its ability to provide the lacking tRNAs corresponding to these *E. coli* rare codons and thus to facilitate expression of heterologous genes [31]. In case of $\Delta 19\text{NCS}$, the strategy to use this strain as expression host significantly improved the expression yield compared to the other strains tested. Concerning the solubility of $\Delta 19\text{NCS}$ expressed in *E. coli* Rosetta(DE3), more than 50% of the recombinant protein was present in the supernatant after cell lysis (Fig. 2, lanes 2 and 4).

Purification of $\Delta 19\text{NCS}$

Purification of soluble $\Delta 19\text{NCS}$ with C-terminal His₆-tag was carried out using nickel ion affinity chromatography. The protein eluted at relatively low imidazole concentrations of 45–60 mM, probably due to bad accessibility of the affinity tag in the folded protein. The corresponding fractions and the flow-through were analyzed by SDS-PAGE (Fig. 2, lanes 6–9). Lanes 7–9 show two intense bands corresponding to molecular masses of about 30 and 24 kDa, respectively. The corresponding proteins are

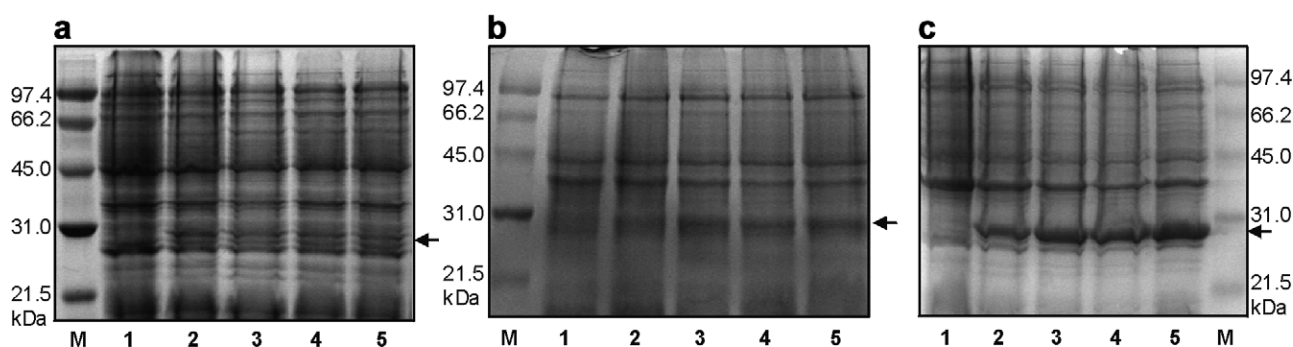


Fig. 1. Comparison of $\Delta 19\text{NCS}$ expression levels in different *E. coli* host strains analyzed by 15% SDS-PAGE. (a) *E. coli* ER2566pLysS/pET29b- $\Delta 19\text{NCS}$; (b) *E. coli* BL21(DE3)/pET29b- $\Delta 19\text{NCS}$; (c) *E. coli* Rosetta(DE3)/pET29b- $\Delta 19\text{NCS}$; M: Molecular weight standard; lanes 1–5: before induction, 1 h, 2 h, 3 h, 4 h after induction; arrow = $\Delta 19\text{NCS}$.

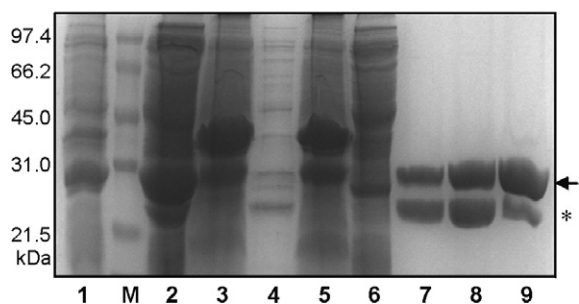


Fig. 2. Purification of $\Delta 19\text{NCS}$ expressed in *E. coli* Rosetta(*DE3*)/pET29b- $\Delta 19\text{NCS}$ grown on M9 minimal medium and analyzed by 15% SDS-PAGE. M: Molecular weight standard; lane 1: yield of $\Delta 19\text{NCS}$ after 4 h expression; lanes 2, 4: soluble fraction after lysis; lanes 3, 5: insoluble fraction after lysis; lanes 6–9: IMAC purification; lane 6: flow-through; lanes 7–9: $\Delta 19\text{NCS}$ eluted at 45–60 mM imidazole; arrow = $\Delta 19\text{NCS}$ full length; * = $\Delta\Delta 19\text{NCS}$.

not present in the flow-through (lane 6). The upper band runs at the same height as the over-expressed protein detected in the cells after induction (lane 1), whereas the lower band is almost invisible at this state, but appears in lanes 2 and 4 analyzing the protein content of the supernatant after cell lysis. As for pure $\Delta 19\text{NCS}$ a single band corresponding to a molecular mass of 25.8 kDa was expected, the two bands were subjected to protein mass fingerprinting in order to identify $\Delta 19\text{NCS}$. The analysis showed that the upper band contained full-length $\Delta 19\text{NCS}$, while the lower band corresponded to a specific degradation product $\Delta\Delta 19\text{NCS}$. The exact N-terminal sequence of the degradation product could be determined by Edman degradation as LHHQG, revealing that the S-tag and 10 additional N-terminal amino acids of $\Delta 19\text{NCS}$ had been cleaved off. In consequence, $\Delta\Delta 19\text{NCS}$ was termed $\Delta 29\text{NCS}$. Though all lysis and purification steps were carried out thoroughly on ice or with ice-cold buffers and in presence of protease inhibitors, specific degradation of $\Delta 19\text{NCS}$ could not be avoided.

In order to separate the S-tag from full-length $\Delta 19\text{NCS}$, thrombin cleavage was performed during dialysis of the dilute protein solution against binding buffer overnight, which turned out to be as efficient as using concentrated protein solution in the special thrombin cleavage buffer containing CaCl_2 recommended by Novagen. Obviously, no protease inhibitor could be used during this step. SDS-PAGE analysis indicates nearly 100% efficiency of thrombin cleavage because no S-tagged $\Delta 19\text{NCS}$ is detectable after cleavage, but an additional band corresponding to the correct 23 kDa thrombin cleavage product as confirmed by protein mass fingerprinting appears (Fig. 3). As the intensity of the lower band corresponding to $\Delta 29\text{NCS}$ increases slightly after cleavage, degradation as described above still takes place during cleavage, but slowly. During the second nickel ion affinity chromatography step performed in order to remove the cleaved S-tag, $\Delta 19\text{NCS}$ and $\Delta 29\text{NCS}$ showed slightly different elution behavior, that is $\Delta 29\text{NCS}$ started to elute at lower imidazole concen-

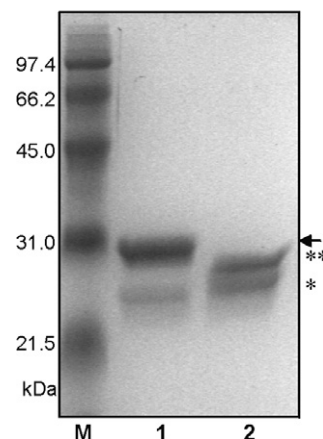


Fig. 3. Thrombin cleavage of S-tag from $\Delta 19\text{NCS}$ analyzed by 15% SDS-PAGE. M: molecular weight standard; lane 1: before thrombin cleavage; lane 2: after thrombin cleavage; arrow: $\Delta 19\text{NCS}$; *: $\Delta\Delta 19\text{NCS}$; **: $\Delta 19\text{NCS}$ without S-tag.

trations compared to $\Delta 19\text{NCS}$. In consequence, the two NCS species could be at least partially separated.

During sample preparation for NMR spectroscopic analysis of the folding state of NCS, samples consisting in major parts of $\Delta 29\text{NCS}$ turned out to be more stable against macroscopic aggregation than samples containing an excess of $\Delta 19\text{NCS}$. NMR spectra of both samples containing mainly $\Delta 19\text{NCS}$ and samples with an excess of $\Delta 29\text{NCS}$ indicated existence of a well defined tertiary structure. Thus, we decided to construct an expression vector coding for $\Delta 29\text{NCS}$ lacking the N-terminal S-tag but including a C-terminal His₆-tag with two central aims. First, we wanted to shorten the purification procedure by avoiding thrombin cleavage of the S-tag not needed for our purposes and, secondly, we wanted to obtain only one defined NCS species suitable for structural NMR spectroscopic studies as well as for quantitative substrate binding studies.

Expression and purification of $\Delta 29\text{NCS}$

The plasmid pET29b- $\Delta 29\text{NCS}$ coding for amino acids 30–210 of NCS linked with a C-terminal His₆-tag was constructed as described above, and the expected sequence was confirmed by DNA sequencing. Prior to preparative expression of labeled protein in M9 medium, expression tests in *E. coli* Rosetta(*DE3*) in LB medium were performed. The expression rate of $\Delta 29\text{NCS}$ was as high as for $\Delta 19\text{NCS}$ in the same host strain. A solubility test comparing levels of soluble NCS after 4, 6 and 20 h induction time showed that an induction period of 6 h is suitable for maximal yields of soluble protein. Preparative expression of $\Delta 29\text{NCS}$ in minimal medium is documented by SDS-PAGE in Fig. 4a. The expression level in M9 medium was as high as in LB medium. Around 3 g cells per liter M9 medium (wet weight) could be harvested after 6 h expression. More than 50% of the overexpressed $\Delta 29\text{NCS}$ was

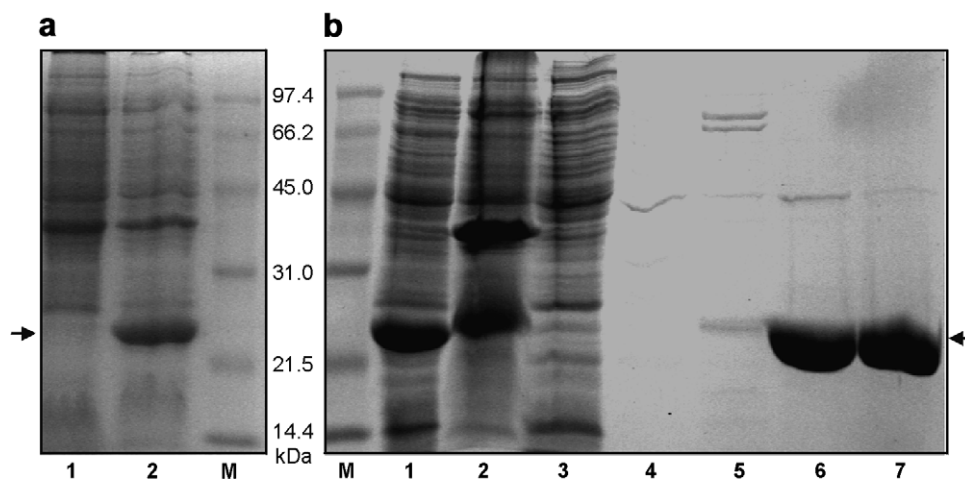


Fig. 4. Expression and purification of $\Delta 29\text{NCS}$ expressed in *E. coli* Rosetta(*DE3*)/pET29b- $\Delta 29\text{NCS}$ grown on M9 minimal medium and analyzed by 15% SDS-PAGE. (a) Expression yield. M: molecular weight standard; lane 1: before induction; lane 2: 6 h after induction; (b) lysis and purification. M: molecular weight standard; lane 1: soluble fraction after lysis; lane 2: insoluble fraction after lysis; lanes 3–7: IMAC purification. Lane 3: flow-through; lanes 4–5: proteins eluted at 20–30 mM imidazole; lanes 6–7: $\Delta 29\text{NCS}$ eluted at 45–60 mM imidazole; arrow: $\Delta 29\text{NCS}$.

found in the soluble fraction after cell lysis (Fig. 4b, lane 1). IMAC was used as single purification step. Highly pure $\Delta 29\text{NCS}$ eluted at imidazole concentrations from 45 to 60 mM (Fig. 4b, lanes 6 and 7). Although the running height of a minor impurity corresponding to a molecular weight of about 45 kDa could be interpreted as a hint on dimerization of $\Delta 29\text{NCS}$, this impurity has also been observed after purification of other proteins using a HisTrap column, and thus it is more likely an abundant *E. coli* protein with weak affinity to the nickel column. As lyophilized protein is easy to store and resistant against degradation, $\Delta 29\text{NCS}$ was dialyzed against H_2O , frozen in liquid nitrogen, lyophilized, and stored at 4 °C until further use. Purification yields ranged from 40 to 50 mg $\Delta 29\text{NCS}$ per liter M9 medium, corresponding to about 15 mg protein per gram cells.

Characterization of $\Delta 29\text{NCS}$ by ESI-MS

Mass spectrometry was used to check identity and purity of recombinant $\Delta 29\text{NCS}$. ESI-MS spectra of recombinant $\Delta 29\text{NCS}$ (Fig. 5) indicate the presence of highly pure species in the sample. Molecular mass of $\Delta 29\text{NCS}$ was determined to be 21179.5 Da which is in agreement

with that calculated from the amino acid sequence (theoretical molecular mass: 21181.3 Da).

Enzymatic activity of $\Delta 29\text{NCS}$

To investigate whether the purified recombinant protein shows enzymatic activity, non-labeled 4-HPAA and radiolabeled dopamine was incubated in presence of native $\Delta 29\text{NCS}$, and, as a negative control, in presence of thermally denatured $\Delta 29\text{NCS}$. Components of the reaction mixture were separated by thin-layer chromatography and radiolabeled reaction products were detected by autoradiography [1] (Fig. 6). In addition to the radiolabeled dopamine visible in the presence of native (lane 1) and denatured (lane 2) $\Delta 29\text{NCS}$, a second radiolabeled reaction product is detectable only in the reaction mixture containing native $\Delta 29\text{NCS}$. The corresponding spot on lane 1 marked with an arrow shows an R_f of 0.60 which is identical to the expected migration distance of the reaction product (*S*)-norcoclaurine [21], showing that the removal of 10 aminoterminal amino acids from the enzymatically active $\Delta 19\text{NCS}$ did not affect enzyme activity. The activity of the recombinant enzyme is further indication for a largely intact native conformation after purification.

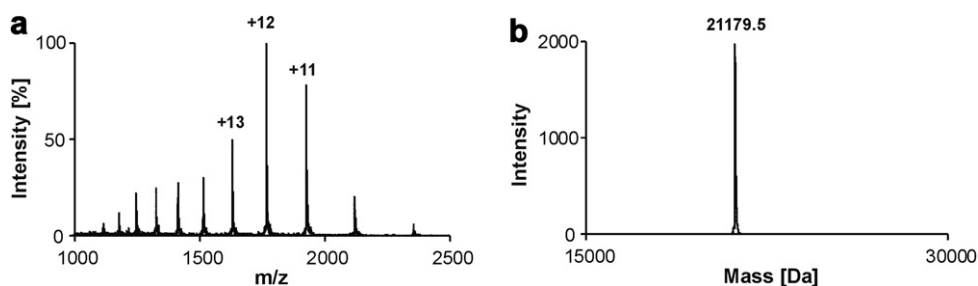


Fig. 5. Positive ion ESI mass spectrum (a) and transformed spectrum (b) of $\Delta 29\text{NCS}$ (100 μM) dissolved in 1 mM ammonium acetate, 5% ethanol and 0.1% acetic acid. Experimental molecular mass of $\Delta 29\text{NCS}$ is 21179.5 Da.

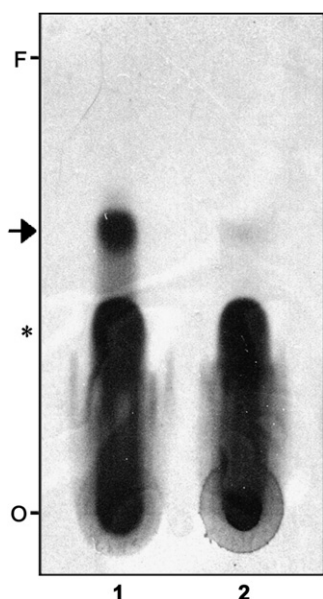


Fig. 6. Enzyme activity assay: autoradiograph of TLC-plate used to separate reaction compounds after 1.5 h incubation at 37 °C. Lane 1: assay containing native $\Delta 29\text{NCS}$; lane 2: negative control containing thermally denatured $\Delta 29\text{NCS}$; arrow: ^{14}C -(S)-norcoclaurine; *: [$8\text{-}^{14}\text{C}$] dopamine; O: sample origin; F: solvent front.

NMR spectroscopy

Proper folding of the purified recombinant $\Delta 29\text{NCS}$ was examined by measuring an ^1H – ^{15}N HSQC spectrum of uniformly ^{15}N -labeled protein. The protein consists of 190 amino acids (Fig. 7a) including nine prolines that do not

a

10	20	30	40	50
MLHHQGIINQ	VSTVTKVIHH	ELEVAASADD	IWTVYSWPGL	AKHLPDLLPG
60	70	80	90	100
AFEKLEIIGD	GGVGTILDMT	FVPGEFPHEY	KEKFILVDNE	HLKKVQMIE
110	120	130	140	150
GGYLDLGVTY	YMDTIHVVT	GKDSCVIKSS	TEYHVKPEFV	KIVEPLITTG
160	170	180	190	
PLAAMADAIS	KLVLEHKSKS	NSDEIEAAII	TVLEHHHHHH	

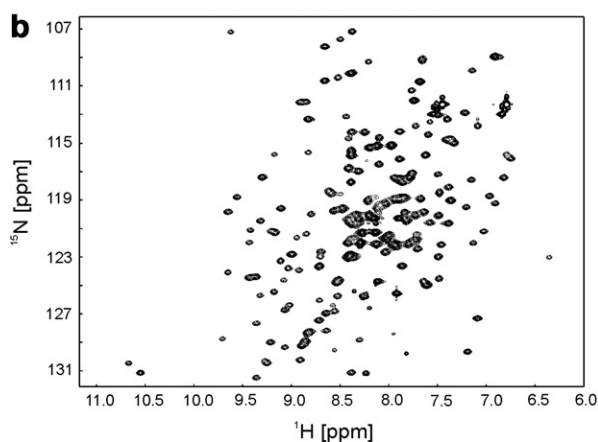


Fig. 7. Analysis of $\Delta 29\text{NCS}$ by NMR spectroscopy. (a) Amino acid sequence of $\Delta 29\text{NCS}$. (b) ^1H – ^{15}N HSQC spectrum of $\Delta 29\text{NCS}$ (730 μM) at 308 K; 800 MHz; 16 accumulations; 256×1024 data points.

contribute signals to the ^1H – ^{15}N HSQC spectrum. In consequence, together with the side chain amide protons, about 200 cross peaks would be expected to show up in the spectrum (Fig. 7b), a number that compares well to the actual number of 194 ^1H – ^{15}N cross correlation peaks detected. Overlapping signals in the spectral region around 8 ppm in the proton dimension and around 121 ppm in the nitrogen dimension indicate the presence of unstructured areas in the protein, but the high dispersion of signals in both dimensions indicates a well defined tertiary structure of $\Delta 29\text{NCS}$. Concluding, the expression and purification protocol described for the newly designed NCS construct $\Delta 29\text{NCS}$ is suitable to yield enough properly folded and enzymatically active protein to allow determination of the three-dimensional solution structure by NMR spectroscopy which is currently in progress in our department.

Acknowledgments

We thank Katrin Weiss for professional technical assistance. This work was supported by Grant DFG RO 617/11-4 from Deutsche Forschungsgemeinschaft.

References

- [1] N. Samanani, D.K. Liscombe, P.J. Facchini, Molecular cloning and characterization of norcoclaurine synthase, an enzyme catalyzing the first committed step in benzyloquinoline alkaloid biosynthesis, *Plant J.* 40 (2004) 302–313.
- [2] D.K. Liscombe, B.P. MacLeod, N. Loukanina, O.I. Nandi, P.J. Facchini, Evidence for the monophyletic evolution of benzyloquinoline alkaloid biosynthesis in angiosperms, *Phytochemistry* 66 (2005) 2501–2520.
- [3] K. Hoffmann-Sommergruber, Plant allergens and pathogenesis-related proteins. What do they have in common? *Int. Arch. Allergy Immunol.* 122 (2000) 155–166.
- [4] H. Breiteneder, C. Ebner, Molecular and biochemical classification of plant-derived food allergens, *J. Allergy. Clin. Immunol.* 106 (2000) 27–36.
- [5] C. Ebner, K. Hoffmann-Sommergruber, H. Breiteneder, Plant food allergens homologous to pathogenesis-related proteins, *Allergy* 56 (Suppl. 67) (2001) 43–44.
- [6] A.B. Kay, Allergy and allergic diseases—first of two parts, *N. Engl. J. Med.* 344 (2001) 30–37.
- [7] A.B. Kay, Allergy and allergic diseases—second of two parts, *N. Engl. J. Med.* 344 (2001) 109–113.
- [8] M. Wills-Karp, J. Santeliz, C.L. Karp, The germless theory of allergic disease: revisiting the hygiene hypothesis, *Nat. Immunol.* 1 (2001) 69–75.
- [9] D. Venarske, R.D. deShazo, Molecular mechanisms of allergic disease, *South Med. J.* 96 (2003) 1049–1054.
- [10] H.J. Gould, B.J. Sutton, A.J. Beavil, R.L. Beavil, N. McCloskey, H.A. Coker, D. Fear, L. Smurthwaite, The biology of IGE and the basis of allergic disease, *Annu. Rev. Immunol.* 21 (2003) 579–628.
- [11] S.M. Miescher, M. Vogel, Molecular aspects of allergy, *Mol. Aspects Med.* 23 (2002) 413–462.
- [12] I.K. Budde, R.C. Aalberse, Histamine-releasing factors, a heterogeneous group of different activities, *Clin. Exp. Allergy* 33 (2003) 1175–1182.
- [13] R. Valenta, S. Vrtala, Recombinant allergens for specific immunotherapy, *Allergy* 54 (Suppl. 56) (1999) 43–44.
- [14] M. Gajhede, P. Osmark, F.M. Poulsen, H. Ipsen, J.N. Larsen, R.J. Joost van Neerven, C. Schou, H. Lowenstein, M.D. Spangfort, X-ray

- and NMR structure of Bet v 1, the origin of birch pollen allergy, *Nat. Struct. Biol.* 3 (1996) 1040–1045.
- [15] C. Faber, A. Lindemann, H. Sticht, A. Ejchart, A. Kungl, M. Susani, R.W. Frank, D. Kraft, M. Breitenbach, P. Rosch, Secondary structure and tertiary fold of the birch pollen allergen Bet v 1 in solution, *J. Biol. Chem.* 271 (1996) 19243–19250.
- [16] P. Neudecker, K. Schweimer, J. Nerkamp, S. Scheurer, S. Vieths, H. Sticht, P. Rosch, Allergic cross-reactivity made visible: solution structure of the major cherry allergen Pru av 1, *J. Biol. Chem.* 276 (2001) 22756–22763.
- [17] P. Neudecker, J. Nerkamp, A. Eisenmann, A. Nourse, T. Lauber, K. Schweimer, K. Lehmann, S. Schwarzing, F. Ferreira, P. Rosch, Solution structure, dynamics, and hydrodynamics of the calcium-bound cross-reactive birch pollen allergen Bet v 4 reveal a canonical monomeric two EF-hand assembly with a regulatory function, *J. Mol. Biol.* 336 (2004) 1141–1157.
- [18] P. Rosch, Nuclear magnetic resonance studies of allergens, *J. Chromatogr. B Biomed. Sci. Appl.* 756 (2001) 165–177.
- [19] M. Saraste, P.R. Sibbald, A. Wittinghofer, The P-loop—a common motif in ATP- and GTP-binding proteins, *Trends Biochem. Sci.* 15 (1990) 430–434.
- [20] P. Neudecker, K. Lehmann, J. Nerkamp, T. Haase, A. Wangorsch, K. Fotisch, S. Hoffmann, P. Rosch, S. Vieths, S. Scheurer, Mutational epitope analysis of Pru av 1 and Api g 1, the major allergens of cherry (*Prunus avium*) and celery (*Apium graveolens*): correlating IgE reactivity with three-dimensional structure, *Biochem. J.* 376 (2003) 97–107.
- [21] N. Samanani, P.J. Facchini, Isolation and partial characterization of norcoclaurine synthase, the first committed step in benzyloquinoline alkaloid biosynthesis, from opium poppy, *Planta* 213 (2001) 898–906.
- [22] T.C. Birdsall, G.S. Kelly, Berberine: therapeutic potential of an alkaloid found in several medicinal plants, *Alt. Med. Rev.* 2 (1997) 94–103.
- [23] M.A. Rashid, K.R. Gustafson, Y. Kashman, J.H. Cardellina, J.B. McMahon, M.R. Boyd, Anti-HIV alkaloids from *Toddalia asiatica*, *Nat. Prod. Lett.* 6 (1995) 153–156.
- [24] J. Sambrook, E.F. Fritsch, M. Maniatis, *Molecular Cloning: a Laboratory Manual*, Cold Spring Harbor Laboratory Press, New York, 1989.
- [25] O. Meyer, H.G. Schlegel, Biology of aerobic carbon monoxide-oxidizing bacteria, *Annu. Rev. Microbiol.* 37 (1983) 277–310.
- [26] U.K. Laemmli, Cleavage of structural proteins during the assembly of the head of bacteriophage T4, *Nature* 227 (1970) 680–685.
- [27] Z. Zhang, A.G. Marshall, A universal algorithm for fast and automated charge state deconvolution of electrospray mass-to-charge ratio spectra, *J. Am. Soc. Mass Spectrom.* 9 (1998) 225–233.
- [28] M.M. Bradford, A rapid and sensitive method for quantitation of microgram quantities of protein using the principle of protein dye binding, *Anal. Biochem.* 72 (1976) 248–254.
- [29] S. Mori, C. Abeygunawardana, M.O. Johnson, P.C. van Zijl, Improved sensitivity of HSQC spectra of exchanging protons at short interscan delays using a new fast HSQC (FHSQC) detection scheme that avoids water saturation, *J. Magn. Reson. B* 108 (1995) 94–98.
- [30] E. Goldman, A.H. Rosenberg, G. Zubay, F.W. Studier, Consecutive low-usage leucine codons block translation only when near the 5' end of a message in *Escherichia coli*, *J. Mol. Biol.* 245 (1995) 467–473.
- [31] J.F. Kane, Effects of rare codon clusters on high-level expression of heterologous proteins in *Escherichia coli*, *Curr. Opin. Biotechnol.* 6 (1995) 494–500.

Einzelarbeit C

Hanna Berkner, Kristian Schweimer, Irena Matečko und Paul Rösch (2008): Conformation, catalytic site, and enzymatic mechanism of the PR10 allergen-related enzyme norcoclaurine synthase. *Biochemical Journal* **413**(2), 281-290

Conformation, catalytic site, and enzymatic mechanism of the PR10 allergen-related enzyme norcoclaurine synthase

Hanna BERKNER, Kristian SCHWEIMER, Irena MATECKO and Paul RÖSCH¹

Research Center for Bio-Macromolecules, Universität Bayreuth Universitätsstrasse 30, 95447 Bayreuth, Germany

The enzyme NCS [(S)-norcoclaurine synthase; EC 4.2.1.78] found in the common meadow rue, *Thalictrum flavum*, and other plant species, is involved in the biosynthesis of BIAs (benzylisoquinoline alkaloids). This group of plant secondary metabolites comprises pharmacologically-active compounds such as morphine and codeine. NCS catalyses the condensation of 4-HPAA (4-hydroxyphenylacetaldehyde) and dopamine to (S)-norcoclaurine, the common precursor of all plant BIAs. Although enzymatic properties of NCS and mechanistic aspects of the reaction have been studied in detail, no structural information on NCS was available so far. The enzyme shows significant sequence homology to members of the PR10 proteins (class 10 of pathogenesis-related proteins) such as the major birch pollen allergen Bet v 1. Our CD and NMR spectroscopic data indicated high similarity of the NCS and the Bet v 1 fold and

allowed us to model NCS using Bet v 1 as a template. Virtually complete backbone assignment of the NCS sequence was used to study substrate binding by NMR titration experiments. Although binding of 4-HPAA seems to induce side-chain rearrangements in an extensive part of the protein, the putative distinct interaction site for dopamine could be clearly identified. The oligomerization state of NCS that reportedly plays an important role in enzyme functionality was determined to be concentration-dependent by SEC (size-exclusion chromatography) as well as NMR relaxation measurements, and the enzyme was found to be predominantly a monomer at the low micromolar concentrations used for activity assays.

Key words: Bet v 1 allergen, homology modelling, norcoclaurine synthase, NMR spectroscopy, substrate binding.

INTRODUCTION

BIAs (benzylisoquinoline alkaloids) form an important group of plant secondary metabolites including approx. 2500 different compounds. Many of these small organic molecules show pharmacological activity and have been used for medical purposes from 3000 years ago to today [1]. BIAs are mainly produced by representatives of the plant families *Papaveraceae*, *Fumariaceae*, *Ranunculaceae*, *Berberidaceae* and *Menispermaceae*. Opium poppy (*Papaver somniferum*) for example produces the analgesics morphine and codeine, the muscle relaxant papaverine, the antibiotic sanguinarine, and noscapine, which is supposed to be anti-tumorigenic [2]. Other less well-known plant species such as the common meadow rue (*Thalictrum flavum*) also produce BIAs with pharmacological activity, in this case the antimicrobial berberine and the potential anti-HIV agent magnoflorine [3,4]. Although humans use BIAs as a pharmaceutical drug, plants probably developed these substances for defence against microbial and herbivore attacks. Thus it is not surprising that the activity of several enzymes involved in BIA biosynthesis is upregulated under environmental stress conditions that can be simulated e.g. by elicitor treatment of cell cultures. This behaviour was demonstrated for native NCS [(S)-norcoclaurine synthase] from *Papaver somniferum* [2]. NCS catalyses the condensation of 4-HPAA (4-hydroxyphenylacetaldehyde) and dopamine to (S)-norcoclaurine (Figure 1), which is the central precursor of all BIAs of plant origin [5,6].

In 2004, the gene coding for NCS in *T. flavum* was identified [7]. The deduced amino acid sequence showed 50–60% homology with members of the PR10 proteins (class 10 of pathogenesis-

related proteins), which includes the Bet v 1 protein family of pollen and food allergens. PR10 proteins also seem to play a role in the plant defence system against pathogens [8], but, despite numerous efforts, their precise function could not be established so far. However, the enzymatic properties of native NCS isolated from a number of different BIA-producing plant species as well as of recombinant NCS from *T. flavum*, *P. somniferum* and *Coptis japonica* have been investigated [2,7,9–11]. Recently, the reaction mechanism of the catalysed Pictet–Spengler reaction could be elucidated for *T. flavum* NCS [12], but still no information was available concerning the substrate-binding sites and the active centre of the enzyme.

The structural characterization of NCS is a prerequisite for the determination of substrate-binding sites and further functional studies on this protein. Recently, we developed a purification protocol yielding high amounts of pure isotopically-labelled *T. flavum* NCS suitable for structural studies by means of NMR spectroscopy [13]. By combining structural information from CD spectroscopy and NMR experiments, including the assignment of $\Delta 29$ NCS backbone resonances, and the amino acid sequence similarity of NCS to the major birch pollen allergen Bet v 1 (Figure 2), we could build a semi-experimental homology model based on the Bet v 1 crystal structure [14]. Moreover, we could use this model to map the regions of the protein affected by binding of the substrate dopamine, and analogues of the second substrate 4-HPAA, determined by NMR titration experiments. Additionally, we determined the oligomerization state of recombinant NCS applying SEC (size-exclusion chromatography) and NMR relaxation measurements. We discuss our results in the context of the enzymatic properties

Abbreviations used: BIA, benzylisoquinoline alkaloid; CSI, chemical shift index; DTT, dithiothreitol; 4-HPAA, 4-hydroxyphenylacetaldehyde; HSQC, heteronuclear single-quantum coherence; NCS, (S)-norcoclaurine synthase; NOE, nuclear Overhauser effect; PR10 proteins, class 10 of pathogenesis-related proteins; SEC, size-exclusion chromatography; TROSY, transverse relaxation optimized spectroscopy.

¹ To whom correspondence should be addressed (email roesch@unibt.de).

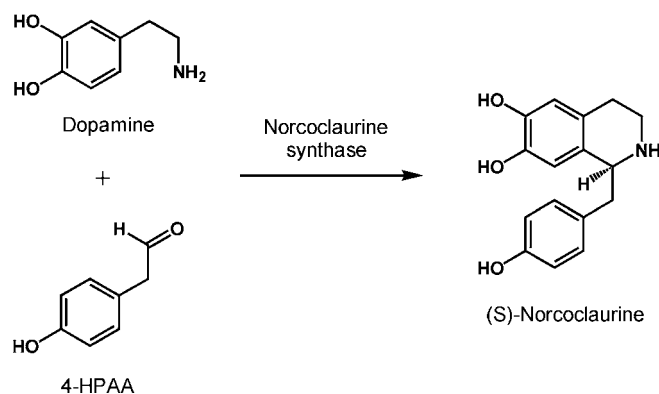


Figure 1 Condensation reaction of 4-HPAA and dopamine catalysed by NCS

This reaction produces (S)-nococlaurine, which is the central precursor of all BIA of plant origin.

$\Delta 29\text{NCS}$	MLHHQGLINOVSTVIRVHHELEVAASADDIWTVYVWPLAKHLPDLPFG	50
Bet v 1	-----GVFNYESVETISVIP-----AARLEKAEILDG-DNLEPKVAPO	36
$\Delta 29\text{NCS}$	AFEKLE--TNGGGVGTILDMTFVPEFPPEYKKEFILLVDNEERLKKVOMI	99
Bet v 1	AISSVENIPGNGGPGTIKKISPEGFPFRVYKDEVDEVDHTNFKYNSVI	86
$\Delta 29\text{NCS}$	EGGYLDLGVTYMMTHVVTGKDSQVIRKSTSEYHVRPEFVKIVEPLITT	149
Bet v 1	EGGPIGDTLEKISNEIKIVATPDGGSILKISNKYHAKGDEHVKAEOVKAS	136
$\Delta 29\text{NCS}$	GPI--AMADAISKVLDPHKSKNSDEIEAAITVLEHHHHHH	190
Bet v 1	KENGETILRAVESYLAHSDAYN-----	159

Figure 2 Sequence alignment of $\Delta 29\text{NCS}$ and the major birch pollen allergen Bet v 1

$\Delta 29\text{NCS}$ and Bet v 1 amino acid sequences were aligned with ClustalW. Black background, identical amino acids; grey background, conservative exchange; and residues enclosed within the box, P-loop consensus sequence. The two homologous proteins show 25% sequence identity.

of NCS reported so far and in comparison with the only enzyme of known structure catalysing a similar Pictet–Spengler-type condensation reaction, strictosidine synthase [15].

EXPERIMENTAL

Materials

All chemicals used in this study were at least 98% pure. Dopamine hydrochloride was purchased from Sigma, whereas L-phenylalanine, methyl(4-hydroxyphenyl)acetate and 2-(4-hydroxyphenyl)ethanol were from Fluka. Isotopically-labelled compounds were purchased from Campro Scientific [Berlin, Germany; ($^{15}\text{NH}_4$) $_2\text{SO}_4$], from Spectra Stable Isotopes (Columbia, MD, U.S.A.; [$\text{U-}^{13}\text{C}$]glucose) or from Euriso-top (Gif-sur-Yvette, France; $^2\text{H}_2\text{O}$).

Expression and purification of $\Delta 29\text{NCS}$

Unlabelled and ^{15}N -labelled $\Delta 29\text{NCS}$ was expressed using the expression system *Escherichia coli* Rosetta(DE3)/pET29b- $\Delta 29\text{NCS}$ as described before [13]. In the case of $^{13}\text{C}^{15}\text{N}$ -labelled $\Delta 29\text{NCS}$, unlabelled glucose was replaced by 2 g/l of [$\text{U-}^{13}\text{C}$]glucose. For expression of deuterated $^{13}\text{C}^{15}\text{N}$ -labelled $\Delta 29\text{NCS}$, five precultures with increasing $^2\text{H}_2\text{O}$ content were grown prior to inoculation of the main culture. The first culture was grown overnight at 37°C in 10 ml of LB (Luria–Bertani)-medium containing 20 $\mu\text{g}/\text{ml}$ kanamycin and 34 $\mu\text{g}/\text{ml}$ chloramphenicol. Subsequent cultures were grown at 37°C in M9 minimal medium [16] including 2 mM MgSO_4 , 10 μM

Fe(III)citrate , 0.1 mM CaCl_2 , 2 g/l [$\text{U-}^{13}\text{C}$]glucose, 1.5 g/l ($^{15}\text{NH}_4$) $_2\text{SO}_4$, trace element solution TS2 [17], 1 \times MEM (minimal essential media) vitamin solution (Gibco), 20 $\mu\text{g}/\text{ml}$ kanamycin and 34 $\mu\text{g}/\text{ml}$ chloramphenicol and contained 25, 50, 75 or 100% $^2\text{H}_2\text{O}$. Each preculture was inoculated to a D_{600} of 0.2 with the respective amount of the preceding preculture. The same medium in 100% $^2\text{H}_2\text{O}$, including 2% of *E. coli* CN Standard medium (^{13}C , ^{15}N labelled; Silantes, Munich, Germany), was used for the final culture with a volume of 1.5 litres. This culture was inoculated to a D_{600} of 0.07 using the whole preceding preculture and incubated at 37°C until it reached a D_{600} of 0.6. After induction by addition of 1 mM IPTG (isopropyl β -D-thiogalactoside), incubation was continued at 25°C and cells were harvested after 9 h of expression. Purification of $\Delta 29\text{NCS}$ was performed as previously described [13]. Pure protein was dialysed against Millipore water, shock-frozen in liquid nitrogen, lyophilized and stored at 4°C.

Analytical SEC

In order to determine the oligomerization state of recombinant $\Delta 29\text{NCS}$, SEC was performed on two sequentially connected Superdex 75 HR 10/30 columns (total bed volume approx. 48 ml; GE Healthcare). Potassium phosphate buffer (50 mM K_2HPO_4), pH 7.0, containing 200 mM NaCl and 1 mM DTT (dithiothreitol) was used as sample buffer as well as for column equilibration and for isocratic protein elution carried out at a flow rate of 0.6 ml/min. Samples with concentrations in a range from 10 μM to 1 mM were prepared either by solubilizing lyophilized $\Delta 29\text{NCS}$ or by dialysing $\Delta 29\text{NCS}$ protein solution against sample buffer. Molecular mass calibration was performed using the Low Molecular Weight (mass) Gel Filtration Calibration Kit (GE Healthcare) comprising the following proteins: albumin (67.0 kDa), ovalbumin (43.0 kDa), chymotrypsinogen (25.0 kDa) and ribonuclease A (13.7 kDa).

CD spectroscopy

Far UV CD measurements were performed on a J-810 S spectropolarimeter with a CDF-426S temperature control unit (JASCO International, Tokyo, Japan). Samples were prepared by dissolving lyophilized protein in 10 mM sodium phosphate buffer (10 mM Na_2HPO_4), pH 7.0, containing 0.5 mM DTT. Spectra were recorded at 20°C and 95°C in a wavelength range of 190–260 nm with a 50 nm/min scanning speed in a 1 mm path length quartz cuvette (Hellma, Mühlheim, Germany) at a protein concentration of 10 μM . Buffer spectra were subtracted and ten spectra were accumulated. Thermal stability was analysed by monitoring the CD signal at 220 nm during heating from 20°C to 95°C and cooling back to 20°C with a heating/cooling rate of 1°C/min. Quartz cuvettes with 1 cm path length equipped with a stirrer were used at a protein concentration of 2.5 μM . In order to normalize the measured ellipticity, the following formula was used to calculate the mean residue molar ellipticity:

$$[\Theta]_{\text{MRW}} = \Theta / (c \cdot d \cdot N)$$

where Θ is measured ellipticity, MRW is the mean residue mass, c is the protein concentration, d is the path length and N is the number of amino acids.

NMR spectroscopy and data evaluation

All NMR experiments were performed at 305.5 K on Bruker Avance 700 MHz and Avance 800 MHz spectrometers with

either room temperature (23 °C) or cryogenically-cooled triple-resonance probes equipped with pulsed field-gradient capabilities.

For determination of longitudinal and transversal ^{15}N relaxation rates (R_1 and R_2) at 700 MHz proton resonance frequency, pulse sequences published by Dayie and Wagner [18] were used. Measurements were performed twice with relaxation delays of 10.6 ms, 531.1 ms, 1062.2 ms, 1593.3 ms and 2124.4 ms for R_1 and 8.8 ms, 26.4 ms, 44.0 ms, 61.6 ms and 79.2 ms for R_2 . Samples with concentrations of 200 μM and 1 mM ^{15}N -labelled $\Delta 29\text{NCS}$ were prepared in 20 mM sodium phosphate buffer, pH 7.8, with 5 mM DTT, 10 % $^2\text{H}_2\text{O}$ and 0.04 % sodium azide. For relaxation measurements in the presence of the substrate dopamine and the 4-HPAA-analogue 2-(4-hydroxyphenyl)ethanol, a 100-fold molar excess of each of the two substances was added to 200 μM ^{15}N -labelled $\Delta 29\text{NCS}$. Relaxation rates were calculated by least-squares fitting of mono-exponential decays to the time-dependent peak intensities with the program Curvefit (<http://biochemistry.hs.columbia.edu/labs/palmer/software/curvefit.html>), and errors were estimated using the jackknife procedure. For estimation of the overall rotational correlation time τ_c with the program Tensor 2.0 [19], residues with relaxation rates differing from the mean value by more than one-fold the standard deviation were excluded. As for globular proteins, τ_c is directly proportional to their dimension, so the molecular mass of a protein with given τ_c can be estimated using a calibration curve calculated from a set of globular proteins of known molecular mass with determined τ_c . We used BMRB (Biological Magnetic Resonance Bank) entries 5841 (14.9 kDa), 4267 (20.7 kDa) and 5720 (23.6 kDa), and the major birch pollen allergen Bet v 1 (17.4 kDa) for calculation of the calibration curve. All τ_c values were normalized to 298 K.

For backbone resonance assignment, the following TROSY (transverse relaxation optimized spectroscopy)-type NMR experiments [20,21] were performed with a 400 μM deuterated ^{15}N , ^{13}C -labelled $\Delta 29\text{NCS}$ sample (in 20 mM sodium phosphate buffer, pH 7.8, 1 mM DTT, 0.04 % sodium azide and 10 % $^2\text{H}_2\text{O}$): [^{15}N , ^1H]-TROSY, tr-HNCO, tr-HN(CA)CO, tr-HNCA, tr-HN(CO)CA, tr-HNCACB and tr-HN(CO)CACB. Additionally, NNH-NOESY and ^{15}N -NOESY-HSQC (^{15}N -NOESY-heteronuclear single-quantum coherence) spectra were recorded with a 670 μM ^{15}N , ^{13}C -labelled $\Delta 29\text{NCS}$ sample in the same buffer. These spectra were principally used to determine the links between the β -strands. NMR data were processed and analysed using in-house written software and the program NMRView 5.2.2 [22]. In order to localize secondary structure elements on the $\Delta 29\text{NCS}$ sequence, the CSI (chemical shift index) was calculated from the chemical shifts of backbone carbonyl and C_α atoms according to Wishart et al. [23]. As perdeuteration of the protein has an influence on chemical shift values [24], the C_α chemical shift was systematically corrected by +0.5 p.p.m. prior to CSI calculation.

All NMR titration experiments were performed with ^{15}N -labelled $\Delta 29\text{NCS}$ in 50 mM sodium phosphate buffer, pH 7.0, containing 2 mM DTT, 0.04 % sodium azide and 10 % $^2\text{H}_2\text{O}$ with two exceptions: first, in the case of the titration with the 4-HPAA-analogue methyl(4-hydroxyphenyl)acetate, for solubility reasons, 5 % DMSO was added to the sample buffer that was also used to prepare the stock solutions for the respective titrations. Secondly, concerning the titration with the 4-HPAA-analogue 2-(4-hydroxyphenyl)ethanol, 20 mM sodium phosphate buffer (pH 7) containing 100 mM NaCl, 0.04 % sodium azide and 10 % $^2\text{H}_2\text{O}$ was used. Substrate binding was investigated by recording a series of [^{15}N , ^1H]-HSQC spectra upon gradual addition of the respective substrate or substrate analogue to a final 100-fold molar excess. The dissociation constant, K_d , was determined by fitting

the ligand concentration-dependent chemical shift changes of residues in the fast exchange limit to the following equation which is valid for a two-state model:

$$\delta_{\text{obs}} = \delta_{\text{P}} + (\delta_{\text{PL}} - \delta_{\text{P}}) \times \left(\frac{K_d + (1+r)P_0}{2P_0} - \frac{[K_d + (1+r)P_0]^2 - 4P_0^2 r}{2P_0} \right)^{1/2}$$

with δ_{obs} , δ_{P} and δ_{PL} as the respective chemical shifts observed in the actual titration step for the free $\Delta 29\text{NCS}$, and for the completely bound ligand. P_0 represents the total concentration of $\Delta 29\text{NCS}$, and r stands for the molar ratio between the respective ligand and $\Delta 29\text{NCS}$. The chemical shift changes between free and bound state in the proton and nitrogen dimension, $\Delta\delta(^1\text{H})$ and $\Delta\delta(^{15}\text{N})$, were normalized to $\Delta\delta_{\text{norm}} = \{\Delta\delta(^1\text{H})^2 + [0.1 \Delta\delta(^{15}\text{N})]^2\}^{1/2}$. Normalized chemical shift changes larger than 0.04 p.p.m. were considered as significant [25].

Sequence alignment and homology modelling

The sequence alignment of $\Delta 29\text{NCS}$ and Bet v 1 was performed using ClustalW [26]. For the three-dimensional structure model, the alignment was modified according to the secondary structure information obtained from the CSI. The $\Delta 29\text{NCS}$ structure model was built with the program Modeller 9v2 [27–29] using Bet v 1 (PDB code 1BV1) as the template structure. The stereochemical quality of the resulting structure was checked with PROCHECK v3.5 [30,31]. Figures were generated using the program PyMOL v0.99 (2006, DeLano Scientific LLC, CA, U.S.A.). Correct orientation of the β -strands was verified according to amide–amide NOEs (nuclear Overhauser effects) observed between the respective backbone amide protons.

RESULTS

Oligomerization state of $\Delta 29\text{NCS}$

Two independent methods, SEC and NMR relaxation measurements, were used to determine the oligomerization state of native recombinant $\Delta 29\text{NCS}$. The apparent molecular mass of the enzyme calculated from both kinds of experiment turned out to be concentration dependent. Most of the measurements were performed two or three times, with the respective S.D. not exceeding 0.4 kDa (Table 1). Although the absolute values for the apparent molecular mass obtained from NMR relaxation experiments are systematically lower than those determined by SEC, they show the same tendency. At concentrations ranging from 10 μM to 1 mM, monomeric and oligomeric forms of $\Delta 29\text{NCS}$ are in fast exchange on the respective timescales, as only the intermediate state can be observed by NMR as well as by SEC. At a concentration of 10 μM , the apparent molecular mass is closer to the molecular mass of the 21.2 kDa monomer, whereas at a concentration of 1 mM it is intermediate between the masses expected for monomeric and dimeric proteins. Thus, $\Delta 29\text{NCS}$ appears to be predominantly monomeric at concentrations around 10 μM . In order to study if dimerization of the enzyme occurs upon substrate binding, measurements in the presence of a 20- or 100-fold excess of the substrate dopamine and the 4-HPAA analogue 2-(4-hydroxyphenyl)ethanol were performed, resulting in absolutely no observable increase in the apparent molecular mass as determined by SEC and only a slight increase of 5 % as determined by NMR relaxation measurements compared with the experiments without substrate. Thus, there is no indication that substrate binding induces dimerization of $\Delta 29\text{NCS}$. The

Table 1 Concentration-dependence of the apparent molecular mass of Δ 29NCS

*Measurements were repeated once or twice; values are given as means \pm S.D. n.d., not detected.

Protein concentration (μ M)	Apparent molecular mass (SEC) (kDa)	Apparent molecular mass (NMR relaxation measurements) (kDa)
10	27.2 \pm 0.1*	n.d.
200	28.8 \pm 0.4*	23.9 \pm 0.2*
400	29.9 \pm 0.2*	n.d.
600	31.6	n.d.
1000	32.5	29.3 \pm 0.2*

difference between the absolute values for the apparent molecular mass calculated from the two methods might result from the fact that the retention volume as well as the relaxation rate is not only dependent on the molecular mass, but also on the hydrodynamic radius of the protein. Even though we report a relatively globular fold for Δ 29NCS and we only chose globular proteins for the calibration curves, the different sets of calibration proteins might at least partially explain this observation.

Secondary structure content and thermal stability of Δ 29NCS versus Bet v 1

CD spectroscopy was used to estimate secondary structure content and thermal stability of Δ 29NCS. The CD spectrum of native Δ 29NCS shows a broad minimum at 215 nm considered characteristic for β -sheet structures (Figure 3A). At 95 °C, this minimum is shifting to lower wavelengths typical for random coil conformation, indicating loss of secondary structure. Subsequent cooling to 20 °C results in a CD spectrum virtually identical to

that of the native protein, showing the reparability of thermal unfolding. The shape of the CD spectra of native Δ 29NCS and native Bet v 1 is highly similar, especially in the minimum region, implying that the two proteins contain similar fractions of the same secondary structure elements (Figure 3B). Secondary structure content of the two proteins was analysed using the CDSSTR algorithm combined with a reference data set consisting of 43 proteins [32] provided by the DICHROWEB server [33,34]. Similar to the secondary structure content calculated for Bet v 1, namely 28 % α -helix and 29 % β -sheet, 27 % of α -helical segments and 25 % β -sheet was calculated for Δ 29NCS.

Temperature-dependent changes of ellipticity at 220 nm (Figure 3C) were used to calculate the melting point of the Δ 29NCS structure as T_m = 66.4 °C, virtually identical to the melting point of Bet v 1, determined to be 67.7 °C. The cooling curve shows thermal unfolding to be only partially reversible, as the ellipticity decreases on cooling, but does not reach the starting point owing to partial protein aggregation.

Structural characterization of Δ 29NCS by NMR spectroscopy

As the concentration-dependent oligomerization of the 21.2 kDa protein leads to increasing transversal relaxation rates, which in turn decrease the signal intensity in 2D (two-dimensional) and especially 3D (three-dimensional) NMR experiments used for backbone assignment, perdeuteration of Δ 29NCS was necessary to yield spectra in a suitable quality for sequential resonance assignment. Nearly all signals visible in the [¹⁵N,¹H]-TROSY spectrum could be assigned (Figure 4A). Most of the unassigned resonances correspond with residual signals of suppressed side-chain amides. Thus 86 % of the Δ 29NCS amino acid sequence could be assigned (Figure 4B). The major part of the signals for the unassigned N-terminal and C-terminal amino acids is

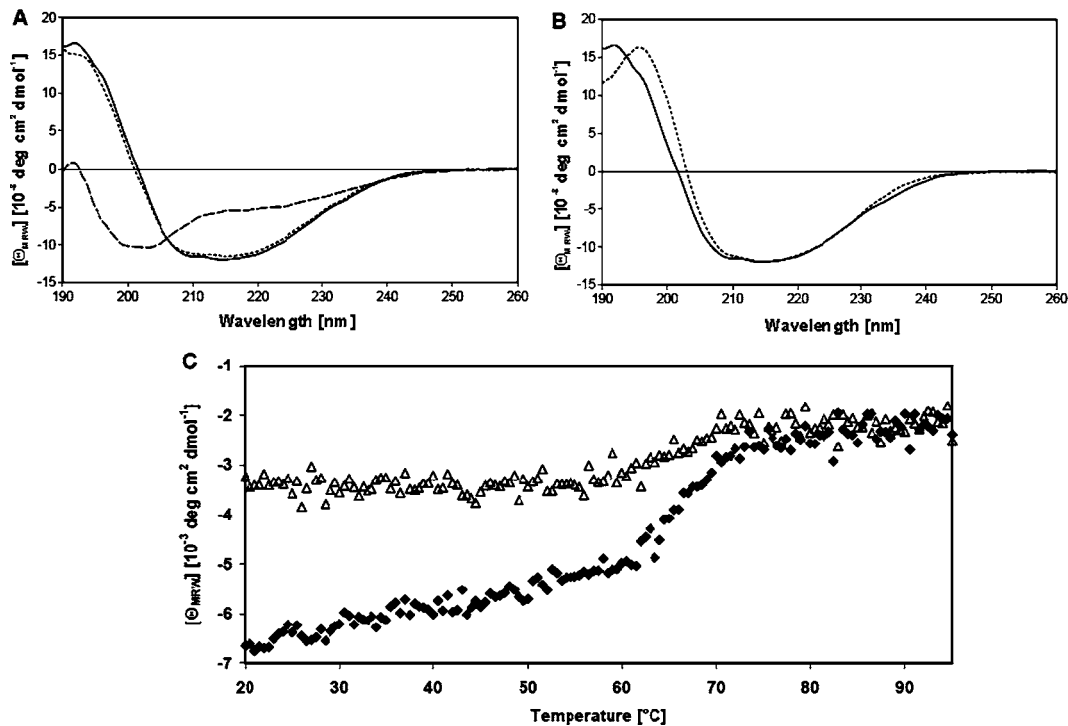


Figure 3 Comparison of Δ 29NCS and Bet v 1 secondary structure content and thermal stability by means of far UV CD spectroscopy

(A) Δ 29NCS in its native state at 20 °C (continuous line), in denatured state at 95 °C (broken line) and renatured at 20 °C (dotted line), calibrated to soluble protein concentration. (B) Native Δ 29NCS (continuous line) and native Bet v 1 (dotted line) at 20 °C. (C) Thermal denaturation of Δ 29NCS: \blacklozenge , heating curve; \triangle , cooling curve.

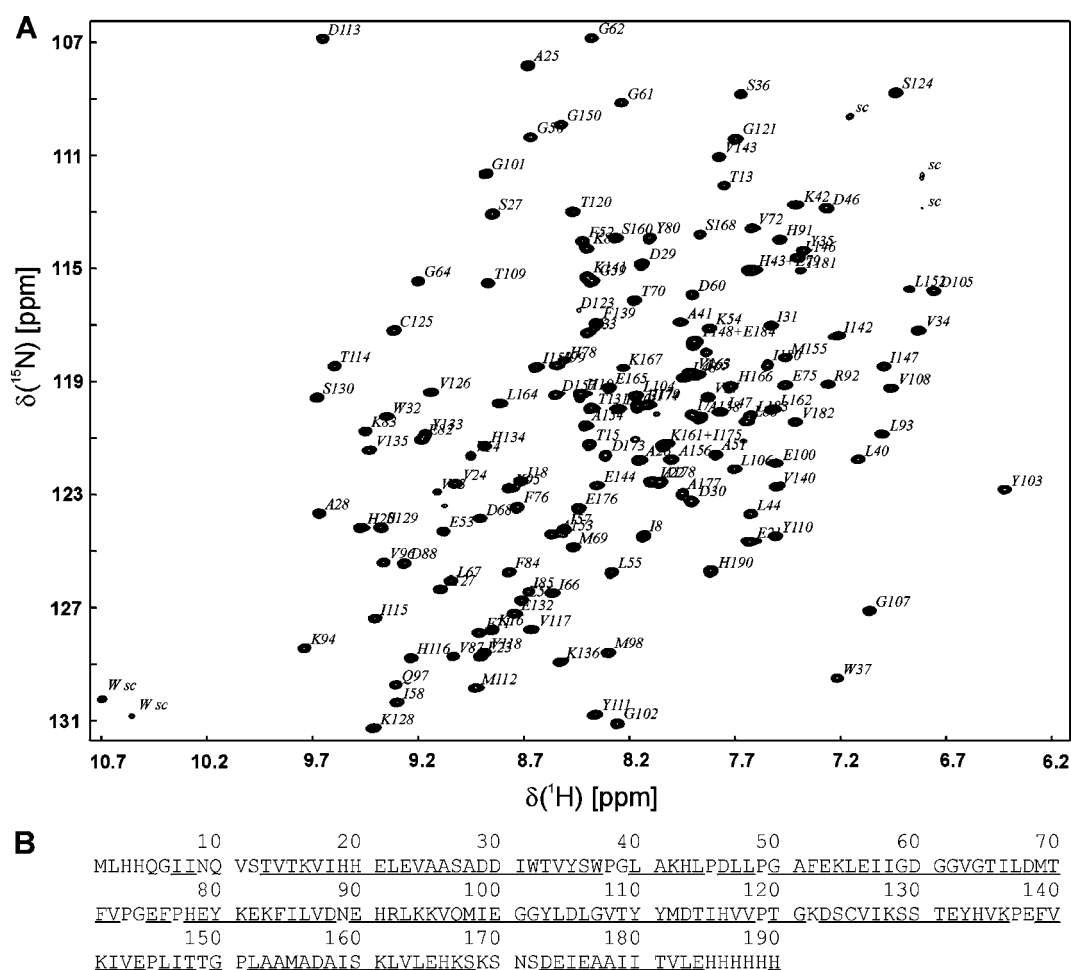


Figure 4 Backbone assignment of $\Delta 29\text{NCS}$

(A) [^{15}N , ^1H]-TROSY NMR spectrum of deuterated ^{15}N , ^{13}C -labelled $\Delta 29\text{NCS}$ (400 μM) measured at 305.5 K and 800 MHz; 4 accumulations; assigned backbone amide resonances are labelled with amino acid type in one letter code and residue number; sc, side-chain amide resonance; W sc, tryptophan side-chain resonance. **(B)** Amino acid sequence of $\Delta 29\text{NCS}$ with assigned residues underlined.

missing owing to conformational or solvent exchange typical for residues in flexible regions and accelerated by the high pH necessarily employed for these measurements. In order to gain information on the distribution of secondary structure elements on the sequence of $\Delta 29\text{NCS}$, we calculated the CSI by combining the C' and C_α chemical shifts. A downfield shift relative to the random coil value, symbolized by positive values in the CSI, indicates the presence of an α -helical conformation, whereas an upfield shift, symbolized by negative values in the CSI, is an indication for β -strands (Figure 5A). The CSI confirms that not only secondary structure content but also distribution of the secondary structure elements is highly similar to Bet v 1, although not entirely identical. As in Bet v 1, the secondary structure aligned to the sequence beginning at the N-terminus starts with a β -strand followed by two α -helices which, in turn, are followed by six β -strands. Major differences between $\Delta 29\text{NCS}$ are found in the C-terminal part; although the major birch pollen allergen contains one long single C-terminal helix, the corresponding helix in $\Delta 29\text{NCS}$ is surrounded by two additional short α -helices. The C-terminal helix is found in the region of the $\Delta 29\text{NCS}$ sequence that cannot be aligned to the shorter Bet v 1. To check whether the identified β -strands are connected to a seven-stranded β -sheet in the same manner as in Bet v 1, we determined their

relative orientation by observing amide–amide NOEs between the backbone amide protons. The NOEs also prove that the relative orientation of the $\Delta 29\text{NCS}$ β -strands is the same as in Bet v 1. This means that in the antiparallel β -sheet, all strands are spatially connected to the respective sequentially succeeding strand, with the exception of the N-terminal β -strand that is orientated antiparallel to the β -strand located most closely to the C-terminus.

Homology model of $\Delta 29\text{NCS}$

The homology model of $\Delta 29\text{NCS}$ was built by combining the experimental results from NMR measurements with homology modelling using the crystal structure of the homologous major birch pollen allergen Bet v 1 [14] as template structure. On the one hand, Bet v 1 was chosen as the template because of sequence similarity, and especially conservation of sequence motifs typical for the Bet v 1 protein family such as the P-loop (Figure 2). On the other hand, it was chosen because the CD data suggested highly similar secondary structure content and thermal stability (Figures 3B and 3C). The CSI (Figure 5A) was used to modify the alignment of $\Delta 29\text{NCS}$ and Bet v 1 obtained by ClustalW according to the expected $\Delta 29\text{NCS}$ secondary

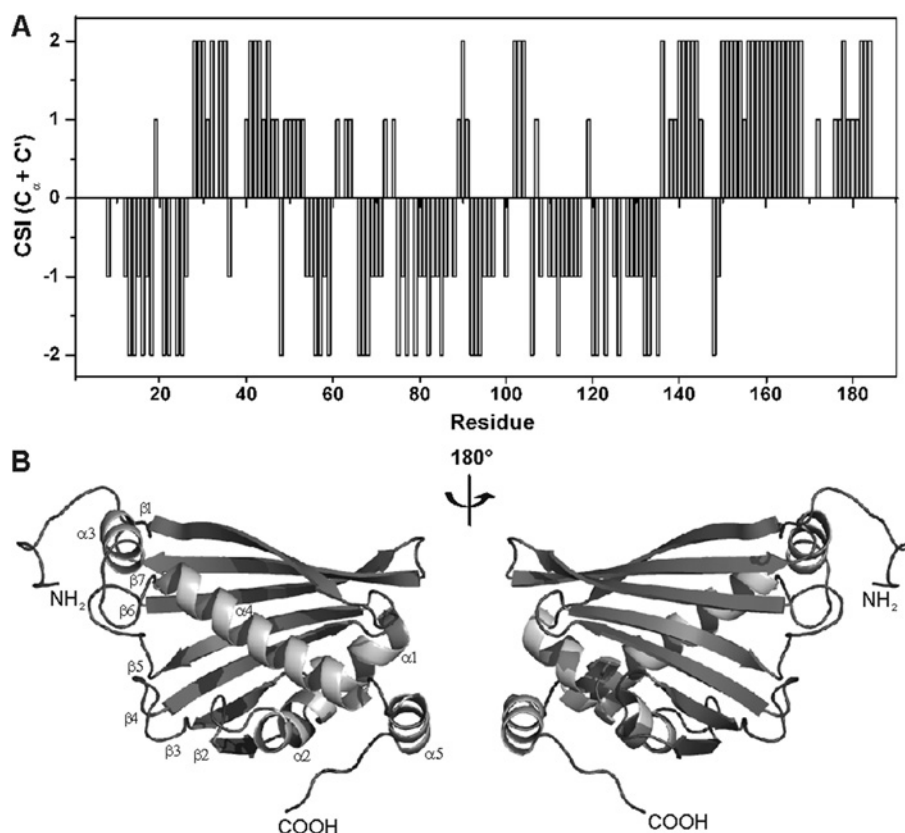


Figure 5 CSI and homology model of $\Delta 29\text{NCS}$

(A) Combined CSI calculated from C' and C_{α} chemical shifts derived from HNCO , HN(CA)CO , HN(CO)CA and HNCA spectra of $\Delta 29\text{NCS}$; +1 was assigned to either C' or C_{α} residues showing a downfield shift related to the random coil value, whereas -1 was assigned to C' or C_{α} residues showing an upfield shift related to the random coil value; 0 was assigned to chemical shifts in the range of values typical for random coil; obtained values for C' and C_{α} were added. (B) Structure of $\Delta 29\text{NCS}$ built by combining homology modelling using the Bet v 1 crystal structure (PDB code 1BV1) as template with experimental information on secondary and tertiary structure as the CSI and backbone amide–amide NOEs in β -strand regions; secondary structure elements are labelled according to their sequence position.

structure. As described before, the CSI suggested that $\Delta 29\text{NCS}$ shows the same secondary structure element distribution as Bet v 1, except for two small α -helices surrounding the long C-terminal α -helix also found in Bet v 1. The locations of these two helices were therefore defined as additional restraints in the modelling procedure. The correct relative orientation of the β -strands in the model structure obtained was verified by comparison to our experimental NOE data. The stereochemical quality of the final model structure (Figure 5B) was assessed using the PROCHECK program. Concerning Φ/Ψ torsion angles, 90% of the amino acid residues were found in the most favoured regions of the Ramachandran plot, whereas 9% and 1% were localized in the additional and generously allowed regions respectively. No residues were found in the disallowed regions. This corresponds to a class 1 stereochemical quality according to Morris et al. [30]. The $\Delta 29\text{NCS}$ model structure consists of an antiparallel seven-stranded β -sheet and two short α -helices, $\alpha 1$ and $\alpha 2$, which pack together against a C-terminal 20-residue α -helix $\alpha 4$, in turn surrounded by two short α -helices, $\alpha 3$ and $\alpha 5$. The latter represent the only secondary structure elements that do not find a counterpart in the Bet v 1 structure. The $\Delta 29\text{NCS}$ β -sheet consists of amino acids 14 \rightarrow 24 ($\beta 1$), 135 \rightarrow 123 ($\beta 7$), 109 \rightarrow 120 ($\beta 6$), 100 \rightarrow 92 ($\beta 5$), 79 \rightarrow 88 ($\beta 4$), 70 \rightarrow 64 ($\beta 3$) and 53 \rightarrow 58 ($\beta 2$). The position of the secondary structure elements is in good agreement with the CSI.

$\Delta 29\text{NCS}$ substrate binding

In order to study substrate binding, NMR titration experiments were performed observing the chemical shift changes of $\Delta 29\text{NCS}$ backbone amide resonances upon gradual addition of either one of the substrates. In the case of 4-HPAA, the substrate analogue methyl(4-hydroxyphenyl)acetate, carrying a methylester group instead of the aldehyde group, was used because 4-HPAA was not available in sufficient purity for quantitative titration experiments. Purity of the oxygen-sensitive compound, which was always handled under oxygen-free conditions, was checked by one-dimensional proton NMR. The addition of methyl(4-hydroxyphenyl)acetate to $\Delta 29\text{NCS}$ provokes significant chemical shift changes for a number of mostly consecutive amino acids (Figure 6A) and, apart from two residues exhibiting chemical shift changes around the significance threshold, the changes in chemical shifts upon addition of dopamine could be related to two sequence areas (Figure 6B).

The amino acids affected in the case of the methyl(4-hydroxyphenyl)acetate titration comprise of residues 37 to 46, corresponding to helix $\alpha 2$; residues 67 to 72, located in the N-terminal part of strand $\beta 3$ and the adjacent loop; residues 81 to 87, representing the major part of strand $\beta 4$; residues 94 to 97, located in the middle of strand $\beta 5$; residues 100 to 103, corresponding to the first part of the adjacent loop; residues 112 to 115, located

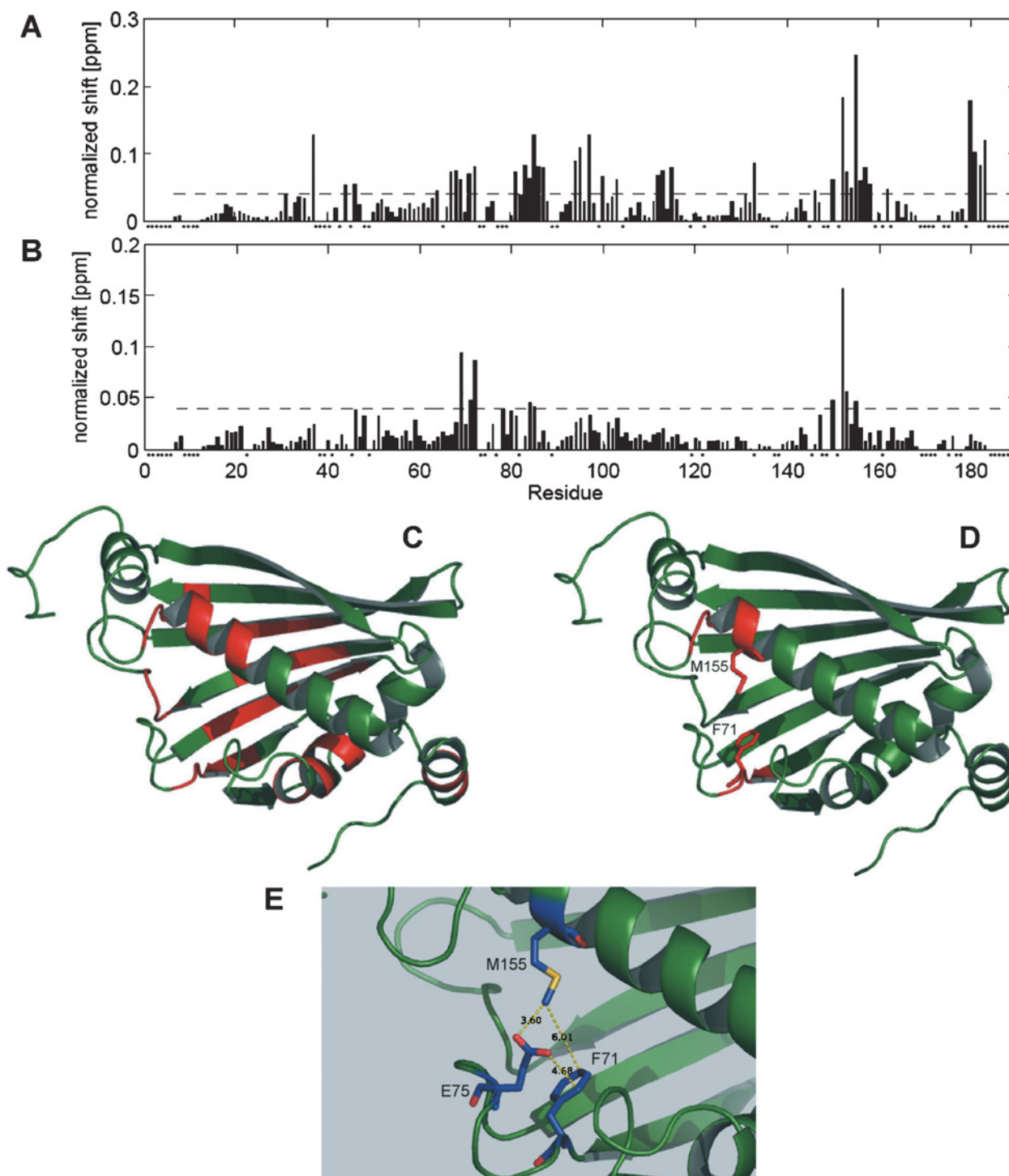


Figure 6 $\Delta 29\text{NCS}$ substrate binding

(A) Chemical shift perturbation expressed in normalized shift derived from the titration of $\Delta 29\text{NCS}$ with up to a 100-fold excess of the 4-HPAA-analogue methyl(4-hydroxyphenyl)-acetate; the broken line represents the significance level of 0.04 p.p.m.; *residue not assigned. (B) Chemical shift perturbation expressed in normalized shift derived from the titration of $\Delta 29\text{NCS}$ with up to a 100-fold excess of dopamine; the broken line represents the significance level of 0.04 p.p.m.; *residue not assigned. (C) $\Delta 29\text{NCS}$ model structure with residues affected by the 4-HPAA analogue binding labelled red. (D) $\Delta 29\text{NCS}$ model structure with residues involved in dopamine binding labelled in red and schematic representation of the potentially involved side chains of Phe⁷¹ and Met¹⁵⁵. (E) Detail of the $\Delta 29\text{NCS}$ model structure showing the orientation of the potential catalytic residue Glu⁷⁵ relative to the suggested binding site of dopamine; distances are in Å (1 Å = 0.1 nm).

in the middle of strand $\beta 6$; residues 150 to 158, representing the N-terminal part of the long C-terminal helix $\alpha 4$; and residues 180 to 183, corresponding to part of helix $\alpha 5$. For helix $\alpha 2$, it is not really obvious from the chemical shift perturbation that a

cluster of amino acids is affected, because the two amino acids subsequent to Trp³⁷ are not assigned, and the chemical shift changes for Leu⁴⁰ and Ala⁴¹ could not be determined as the corresponding signals disappear during titration, indicating

that these residues show intermediate exchange, in contrast to all other residues whose exchange behaviour is fast on the NMR time scale. The only residue showing a significant chemical shift change that is not integrated in a cluster of shifting residues is Tyr¹³³, which is located on strand β 7 and is spatially close to the affected parts of helix α 4 and strand β 6 (Figure 6C). A direct interaction with the relatively small substrate molecule cannot explain the chemical shift changes in such a large portion of the protein. Thus, it is likely that binding of methyl(4-hydroxyphenyl)acetate induces a rearrangement of the side chains in the affected parts of the protein that are all in spatial proximity to each other.

We also used the 4-HPAA-analogue 2-(4-hydroxyphenyl)ethanol for titration, albeit under slightly different buffer conditions. In most cases, the chemical shift changes observed are assigned to the same residues shifting upon addition of methyl(4-hydroxyphenyl)acetate, indicating that the methylester or hydroxy group replacing the aldehyde group in the respective analogue does not alter substrate binding.

In the case of dopamine binding, the chemical shift changes can be found mainly in a distinct part of the protein comprising amino acids 69 to 72 that represent part of strand β 3 and the adjacent loop. Furthermore, residues 150 to 155, corresponding to the N-terminal part of helix α 4, are involved (Figure 6D). In this case, the chemical shift changes are probably induced by direct binding to the respective residues. Closer inspection of the type of amino acids involved and the spatial orientation of their side chains reveals a putative binding site between Phe⁷¹, which might be responsible for hydrophobic interactions with the aromatic ring of dopamine, and Met¹⁵⁵, which might associate with the aliphatic part. A K_d value of 5 mM could be estimated by averaging the values determined from the concentration-dependent chemical shift changes for residues Phe⁷¹ and Met¹⁵⁵, indicating rather weak binding.

To rule out the possibility that we observed non-specific substrate binding, we performed a negative control titration with phenylalanine, yielding chemical shift changes of 0.04 and 0.06 p.p.m. that is in the range of the significance level for only two residues, namely Val⁷² and Leu¹⁵². These residues are also affected by dopamine binding, but the chemical shift changes observed in this case are more than twice as high, and in contrast to phenylalanine binding, the saturation level is almost reached for these residues. Thus, despite the rather high K_d values, it is unlikely that we observe non-specific substrate binding.

Samanani and Facchini [9] conclude from product inhibition experiments that 4-HPAA binds to the enzyme before dopamine. Thus, we also performed titration experiments with dopamine in the presence of one of the 4-HPAA analogues at saturation level. Significant chemical shift changes could not be observed, except for the amino acids Ala¹⁵³ and Met¹⁵⁵, and determination of K_d values was hindered by the fact that we could not reach saturation level in our experiments. This is surprising at first glance, because it could be interpreted as an indication that dopamine does not bind to Δ 29NCS subsequent to 4-HPAA analogue binding. However, as the putative dopamine binding site is located in the part of the protein also affected by 4-HPAA analogue binding, it is likely that the 4-HPAA binding site is located close to the dopamine binding site. Thus, only small changes might occur because the chemical environment is already changed by 4-HPAA analogue binding.

DISCUSSION

The plant enzyme NCS catalyses the condensation of 4-HPAA and dopamine to the central BIA precursor (S)-norcoclaurine [5,6]. So

far, enzymatic properties of NCS from a number of different plant species have been investigated [2,9,35,36] and, in a few cases, the production of recombinant protein could be established [7,10,11]. The oligomerization state that plays an important role for enzyme functionality was investigated for NCS isolated from *T. flavum*. The results from SEC indicated that the enzyme is a dimer, as the molecular mass of NCS derived from SDS/PAGE was determined to be 15 kDa and that determined by SEC was approx. 28 kDa [9]. However, at that time the protein sequence of NCS was unknown. The identification of the gene coding for NCS revealed that the protein consists of 210 amino acids including a potential N-terminal signal peptide of 19 amino acids that is probably cleaved during the subcellular localization process in the plant [7]. Consequently, one would expect a \sim 21 kDa protein to be present in the plant. Therefore, we studied oligomerization of recombinant *T. flavum* Δ 29NCS by two independent methods, namely SEC and NMR relaxation measurements. The results of both methods demonstrated concentration-dependent oligomerization of Δ 29NCS with an equilibrium of monomeric and oligomeric protein over a broad concentration range (Table 1). Although the total values for the apparent molecular mass of Δ 29NCS differed between the two methods, it can be concluded that at low protein concentrations in the range of 10 μ M, the enzyme is predominantly monomeric. Since NCS activity assays were performed at concentrations of approx. 80 nM [7,12], one would expect that the prevailing monomeric form of the enzyme is in large parts responsible for the observed activity. Substrate saturation kinetics measured for natural as well as for recombinant *T. flavum* NCS indicated co-operativity in the case of dopamine binding [7,9,12]. This observation was explained by the assumption that NCS exists as a homodimer with one dopamine binding site on each subunit. It was proposed that binding to one site could influence binding to the site on the other subunit. Our results on the Δ 29NCS oligomerization state do not support this explanation. Furthermore, our experimental titration curves for dopamine binding fit well to a two-state model, suggesting non-co-operative binding. Moreover, we did not find evidence for a second dopamine binding site in the enzyme. Rueffer et al. [35] and Schuhmacher et al. [36] reported non-co-operative substrate saturation kinetics with regard to dopamine for NCS from a number of different plant species, among them the closely-related *Thalictrum tuberosum*.

General enzymatic properties such as pH and temperature optimum were also determined for natural as well as for recombinant *T. flavum* NCS [7,9]. The relatively high temperature optimum at 45–55 °C for the natural and at 40 °C for the recombinant enzyme, and the rapid decrease in activity at temperatures exceeding 55 °C, is in agreement with our results on the thermal stability of Δ 29NCS. The melting point of 66.4 °C corresponds to co-operative unfolding of the protein in the temperature range between 60 °C and 70 °C (Figure 3C).

Whereas the reaction mechanism of the catalysed Pictet–Spengler reaction was established for NCS from *T. flavum* [12], no structural information on NCS has been available. However, this is a prerequisite for the localization of substrate-binding sites. Combining experimental NMR spectroscopic data with homology modelling using the homologous major birch pollen allergen Bet v 1 as template, we have built a reliable model structure of Δ 29NCS (Figure 5B). Since we achieved almost complete backbone assignment of the protein (Figure 4), we could use this model to investigate substrate binding by NMR titration experiments. Using different analogues of 4-HPAA, we observed chemical shift changes for a considerable part of the protein, suggesting that a rearrangement of the respective side chains is induced by 4-HPAA binding (Figures 6A and 6C). This

observation is in agreement with results from product analogue inhibition experiments reported by Samanani and Facchini [9], who suggested that 4-HPAA binds to a different conformation of the enzyme compared with the product (S)-norcoclaurine. They proposed the changes occur upon product release, whereas our data suggests that they are induced by 4-HPAA binding. From the present paper, the exact binding site of 4-HPAA is difficult to determine because the extent of chemical shift changes does not allow us to distinguish between indirectly caused rearrangement and direct interaction. The chemical shift changes observed upon dopamine binding can be assigned to a distinct area in the protein structure, and spatial orientation of the side chains as well as lower K_d values relative to the surrounding amino acids suggest direct interactions of dopamine with amino acid residues Phe⁷¹ and Met¹⁵⁵ (Figures 6B and 6D). The involvement of Phe⁷¹ is supported by the fact that binding pockets of other proteins that bind dopamine also comprise phenylalanine residues [37,38].

The K_d values estimated for dopamine binding are in the mM range, suggesting only weak substrate binding. Since it was suggested that dopamine binds after 4-HPAA [9], the side-chain rearrangement induced by 4-HPAA binding might increase the affinity of the enzyme for dopamine. In addition, the low binding affinities of NCS could be due to its subcellular location. As NCS from *T. flavum* carries a putative signal peptide [7], it is probably targeted to the endoplasmic reticulum. From there, it might be translocated to different cellular compartments such as special vacuoles. Interestingly, it has been shown that *Papaver bracteatum*, which is a BIA-producing plant, accumulates dopamine in vacuolar compartments [39]. In consequence, the high local substrate concentration could compensate for the low binding affinity. The efficiency of the enzyme might thus be due to its stereospecificity with respect to product formation.

The only enzyme of known structure that catalyses a similar Pictet–Spengler-type reaction is strictosidine synthase which, among other plants, was isolated from *Rauvolfia serpentina* [15]. This enzyme catalyses the condensation of the aldehyde secologanine and the amine tryptamine to strictosidine, a reaction analogous to that catalysed by NCS, but in this case yielding the precursor for all plant indole alkaloids. No sequence homology is found between the two enzymes, and the protein fold is also quite different, but there are similarities in catalytic mechanism. In the active site of strictosidine synthase, a glutamate is found that is suggested to deprotonate the amine moiety of tryptamine. This facilitates the formation of a Schiff base with the aldehyde moiety of secologanine. In the $\Delta 29$ NCS structure, Glu⁷⁵ is found in the immediate vicinity of the putative dopamine binding site, with its side chain pointing towards the binding site (Figure 6E). Consequently, this residue could play the same role in norcoclaurine formation, facilitating the Schiff base formation between the deprotonated dopamine and 4-HPAA. Considering the reaction mechanism proposed by Luk et al. [12], it is even more likely that Glu⁷⁵ plays a role in deprotonation of the σ -intermediate formed after the electrophilic cyclization between the iminium ion and the aromatic ring of dopamine.

We thank Ramona Heissmann for expert technical assistance. We would also like to give our thanks to Peter J. Facchini (Department of Biological Sciences, University of Calgary, Calgary, Alberta, Canada T2N 1N4) for providing the original TNCSD $\Delta 19$ /pET29b plasmid. This project was supported by a grant from the Deutsche Forschungsgemeinschaft (RO 617/11-4).

REFERENCES

- Buchanan, B., Gruissem, W. and Jones, R. E. (2000) Biochemistry and Molecular Biology of Plants, American Society of Plant Physiologists, Rockville, MD, U.S.A.
- Samanani, N. and Facchini, P. J. (2001) Isolation and partial characterization of norcoclaurine synthase, the first committed step in benzyloquinoline alkaloid biosynthesis, from opium poppy. *Planta* **213**, 898–906
- Birdsall, T. C. and Kelly, G. S. (1997) Berberine: therapeutic potential of an alkaloid found in several medicinal plants. *Alt. Med. Rev.* **2**, 94–103
- Rashid, M. A., Gustafson, K. R., Kashman, Y., Cardellina, J. H., McMahon, J. B. and Boyd, M. R. (1995) Anti-HIV alkaloids from *Toddalia asiatica*. *Nat. Prod. Lett.* **6**, 153–156
- Stadler, R., Kutchan, T. M., Löffler, S., Nagakura, N., Cassels, B. K. and Zenk, M. H. (1987) Revision of the early steps of reticuline biosynthesis. *Tetrahedron Lett.* **28**, 1251–1254
- Stadler, R., Kutchan, T. M. and Zenk, M. H. (1989) Norcoclaurine is the central intermediate in benzyloquinoline alkaloid biosynthesis. *Phytochemistry* **28**, 1083–1086
- Samanani, N., Liscombe, D. K. and Facchini, P. J. (2004) Molecular cloning and characterization of norcoclaurine synthase, an enzyme catalyzing the first committed step in benzyloquinoline alkaloid biosynthesis. *Plant J.* **40**, 302–313
- Ebner, C., Hoffmann-Sommergruber, K. and Breiteneder, H. (2001) Plant food allergens homologous to pathogenesis-related proteins. *Allergy* **56**, 43–44
- Samanani, N. and Facchini, P. J. (2002) Purification and characterization of norcoclaurine synthase. The first committed enzyme in benzyloquinoline alkaloid biosynthesis in plants. *J. Biol. Chem.* **277**, 33878–33883
- Liscombe, D. K., MacLeod, B. P., Loukanina, N., Nandi, O. I. and Facchini, P. J. (2005) Evidence for the monophyletic evolution of benzyloquinoline alkaloid biosynthesis in angiosperms. *Phytochemistry* **66**, 2501–2520
- Minami, H., Dubouzet, E., Iwasa, K. and Sato, F. (2007) Functional analysis of norcoclaurine synthase in *Coptis japonica*. *J. Biol. Chem.* **282**, 6274–6282
- Luk, L. Y., Bunn, S., Liscombe, D. K., Facchini, P. J. and Tanner, M. E. (2007) Mechanistic studies on norcoclaurine synthase of benzyloquinoline alkaloid biosynthesis: an enzymatic Pictet–Spengler reaction. *Biochemistry* **46**, 10153–10161
- Berkner, H., Engelhorn, J., Liscombe, D. K., Schweimer, K., Wöhrli, B. M., Facchini, P. J., Rösch, P. and Matecko, I. (2007) High-yield expression and purification of isotopically labeled norcoclaurine synthase, a Bet v 1-homologous enzyme, from *Thalictrum flavum* for NMR studies. *Protein Expr. Purif.* **56**, 197–204
- Gajhede, M., Osmark, P., Poulsen, F. M., Ipsen, H., Larsen, J. N., Joost van Neerven, R. J., Schou, C., Lowenstein, H. and Spangfort, M. D. (1996) X-ray and NMR structure of Bet v 1, the origin of birch pollen allergy. *Nat. Struct. Biol.* **3**, 1040–1045
- Ma, X., Panjikar, S., Koepke, J., Loris, E. and Stockigt, J. (2006) The structure of *Rauvolfia serpentina* strictosidine synthase is a novel six-bladed beta-propeller fold in plant proteins. *Plant Cell* **18**, 907–920
- Sambrook, J., Fritsch, E. F. and Maniatis, M. (1989) Molecular Cloning: A Laboratory Manual. Cold Spring Harbor Laboratory Press, New York
- Meyer, O. and Schlegel, H. G. (1983) Biology of aerobic carbon monoxide-oxidizing bacteria. *Annu. Rev. Microbiol.* **37**, 277–310
- Dayie, K. T. and Wagner, G. (1994) Relaxation-rate measurements for ¹⁵N-¹H groups with pulsed-field gradients and preservation of coherence pathways. *J. Magn. Reson. A* **111**, 121–126
- Dosset, P., Hus, J. C., Blackledge, M. and Marion, D. (2000) Efficient analysis of macromolecular rotational diffusion from heteronuclear relaxation data. *J. Biomol. NMR* **16**, 23–28
- Salzmann, M., Pervushin, K., Wider, G., Senn, H. and Wüthrich, K. (1998) TROSY in triple-resonance experiments: new perspectives for sequential NMR assignment of large proteins. *Proc. Natl. Acad. Sci. U.S.A.* **95**, 13585–13590
- Pervushin, K., Riek, R., Wider, G. and Wüthrich, K. (1997) Attenuated T2 relaxation by mutual cancellation of dipole-dipole coupling and chemical shift anisotropy indicates an avenue to NMR structures of very large biological macromolecules in solution. *Proc. Natl. Acad. Sci. U.S.A.* **94**, 12366–12371
- Johnson, B. A. (2004) Using NMRView to visualize and analyze the NMR spectra of macromolecules. *Methods Mol. Biol.* **278**, 313–352
- Wishart, D. S., Sykes, B. D. and Richards, F. M. (1991) Relationship between nuclear magnetic resonance chemical shift and protein secondary structure. *J. Mol. Biol.* **222**, 311–333
- Gardner, K. H. and Kay, L. E. (1998) The use of ²H, ¹³C, ¹⁵N multidimensional NMR to study the structure and dynamics of proteins. *Annu. Rev. Biophys. Biomol. Struct.* **27**, 357–406
- Hajduk, P. J., Dinges, J., Miknis, G. F., Merlock, M., Middleton, T., Kempf, D. J., Egan, D. A., Walter, K. A., Robins, T. S., Shuker, S. B. et al. (1997) NMR-based discovery of lead inhibitors that block DNA binding of the human papillomavirus E2 protein. *J. Med. Chem.* **40**, 3144–3150
- Chenna, R., Sugawara, H., Koike, T., Lopez, R., Gibson, T. J., Higgins, D. G. and Thompson, J. D. (2003) Multiple sequence alignment with the Clustal series of programs. *Nucleic Acids Res.* **31**, 3497–3500

- 27 Marti-Renom, M. A., Stuart, A. C., Fiser, A., Sanchez, R., Melo, F. and Sali, A. (2000) Comparative protein structure modeling of genes and genomes. *Annu. Rev. Biophys. Biomol. Struct.* **29**, 291–325
- 28 Sali, A. and Blundell, T. L. (1993) Comparative protein modelling by satisfaction of spatial restraints. *J. Mol. Biol.* **234**, 779–815
- 29 Fiser, A., Do, R. K. and Sali, A. (2000) Modeling of loops in protein structures. *Protein Sci.* **9**, 1753–1773
- 30 Morris, A. L., MacArthur, M. W., Hutchinson, E. G. and Thornton, J. M. (1992) Stereochemical quality of protein structure coordinates. *Proteins* **12**, 345–364
- 31 Laskowski, R. A., Moss, D. S. and Thornton, J. M. (1993) PROCHECK: a program to check the stereochemical quality of protein structures. *J. Appl. Cryst.* **26**, 283–291
- 32 Sreerama, N. and Woody, R. W. (2000) Estimation of protein secondary structure from circular dichroism spectra: comparison of CONTIN, SELCON, and CDSSTR methods with an expanded reference set. *Anal. Biochem.* **287**, 252–260
- 33 Lobley, A., Whitmore, L. and Wallace, B. A. (2002) DICHROWEB: an interactive website for the analysis of protein secondary structure from circular dichroism spectra. *Bioinformatics* **18**, 211–212
- 34 Whitmore, L. and Wallace, B. A. (2004) DICHROWEB, an online server for protein secondary structure analyses from circular dichroism spectroscopic data. *Nucleic Acids Res.* **32**, W668–W673
- 35 Rueffer, M., El-Shagi, H., Nagakura, N. and Zenk, M. H. (1981) (S)-norlaudanosoline synthase: the first enzyme in the benzyloquinoline biosynthetic pathway. *FEBS Lett.* **129**, 5–9
- 36 Schumacher, H. M., Ruffer, M., Nagakura, N. and Zenk, M. H. (1983) Partial purification and properties of (S)-norlaudanosoline synthase from *Eschscholtzia tenuifolia* cell cultures. *Planta Med.* **48**, 212–220
- 37 Lu, J. H., Li, H. T., Liu, M. C., Zhang, J. P., Li, M., An, X. M. and Chang, W. R. (2005) Crystal structure of human sulfotransferase SULT1A3 in complex with dopamine and 3'-phosphoadenosine 5'-phosphate. *Biochem. Biophys. Res. Commun.* **335**, 417–423
- 38 Erlandsen, H., Flatmark, T., Stevens, R. C. and Hough, E. (1998) Crystallographic analysis of the human phenylalanine hydroxylase catalytic domain with bound catechol inhibitors at 2.0 Å resolution. *Biochemistry* **37**, 15638–15646
- 39 Kutchan, T. M., Rush, M. and Coscia, C. J. (1986) Subcellular localization of alkaloids and dopamine in different vacuolar compartments of *Papaver bracteatum*. *Plant Physiol.* **81**, 161–166

Received 5 February 2008; accepted 26 March 2008

Published as BJ Immediate Publication 26 March 2008, doi:10.1042/BJ20080306

9 Danksagung

An dieser Stelle möchte ich allen danken, die auf direktem oder auch indirektem Wege zum Gelingen dieser Arbeit beigetragen haben:

- Herrn Prof. Dr. Paul Rösch, unter dessen Leitung die vorliegende Arbeit von November 2005 bis Januar 2009 am Lehrstuhl für Struktur und Chemie der Biopolymere der Universität Bayreuth angefertigt wurde. Er hat nicht nur das interessante Thema angeregt, sondern auch für hervorragende Arbeitsbedingungen am Lehrstuhl gesorgt und somit das Gelingen dieser Arbeit in jeglicher Hinsicht gefördert.
- Frau Dr. Irena Matečko für ihre Unterstützung und Hilfsbereitschaft bei der vorliegenden Arbeit.
- Frau Prof. Dr. Birgitta Wöhrl und Herrn Dr. Kristian Schweimer, die jederzeit für Fragen offen waren und mit Vorschlägen und Anregungen zum Gelingen dieser Arbeit beigetragen haben.
- Herrn Prof. Dr. Stefan Vieths (Paul-Ehrlich-Institut, Langen) und Herrn Dr. Jonas Lidholm (Phadia, Uppsala, Schweden) für die freundliche, offene und effektive Zusammenarbeit bezüglich immunologischer Untersuchungen.
- Herrn Prof. Dr. Rob C. Aalberse (Universität Amsterdam, Niederlande) für eine sehr interessante und anregende Diskussion während der ISMA 2007 in Rom.
- Herrn Prof. Dr. Peter J. Facchini (University of Calgary, Kanada) für die Bereitstellung des *Tf*-NCS-Originalplasmids und die Kooperationsbereitschaft.
- Herrn Prof. Dr. Holger Dobbek und Frau Dr. Berta Martins (Bioanorganische Chemie, Universität Bayreuth) für die Möglichkeit zur Nutzung der *Glovebox* und die Betreuung der kristallographischen Versuche.
- Navina Chrobok, die mich im Rahmen ihrer Bachelorarbeit unterstützt und großes Interesse am Thema der Arbeit gezeigt hat.

- Meinen lieben Kollegen innerhalb und außerhalb des Labors, die alle auf ihre Art und Weise zu einer meist ausgesprochen angenehmen und häufig sehr humorvollen Arbeitsatmosphäre beigetragen haben. Insbesondere möchte ich mich bei Ramona Heissmann für technische Unterstützung auf höchstem Niveau, bei Rainer Hofmann für schnelle Hilfe bei Computerproblemen und bei Gudrun Wagner und Violaine Zigan für Unterstützung in Verwaltungsangelegenheiten bedanken.
- Silvia Berkner, Martin Klose, Eva Krammer, Wolfgang Schöpf und meinen Eltern für ausdauernde Unterstützung während der Anfertigung dieser Arbeit und Ratschläge oder aufmunternde Worte im richtigen Moment.

10 Erklärung

Hiermit erkläre ich, dass ich die vorliegende Arbeit selbständig verfasst und keine als die von mir angegebenen Quellen und Hilfsmittel verwendet habe.

Ferner erkläre ich, dass ich nicht anderweitig mit oder ohne Erfolg versucht habe eine Dissertation einzureichen oder mich der Doktorprüfung zu unterziehen.

Bayreuth, den 04.02.2009